Univerzita Karlova v Praze
Fakulta sociálních věd
Institut ekonomických studií

Bakalářská práce

Praha 2011 **Jan Jelínek**

Charles University in Prague
Faculty of Social Sciences
Institute of Economic Studies

Bachelor Thesis

Determinants of Housing Prices in Central and Eastern Europe

Author: Jan Jelínek

Supervisor: PhDr. Michal Hlaváček, PhD.

Academic Year: 2010/2011

Poděkování

Chtěl bych tímto poděkovat PhDr. Michalu Hlaváčkovi, PhD., za cenné rady a podněty při psaní této práce. Dále děkuji PhDr. Ladislavu Krištoufkovi za pomoc v ekonometrické části.

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně a použil pouze uvedené prameny a literaturu. Práce nebyla využita k získání stejného nebo jiného titulu.

V Praze dne

Jan Jelínek

Bibliografický záznam

JELÍNEK, Jan. Determinants of Housing Prices in Central and Eastern Europe. Praha, 2011. 53 s. Bakalářská práce (Bc.) Univerzita Karlova, Fakulta sociálních věd, Institut ekonomických studií. Vedoucí bakalářské práce: PhDr. Michal Hlaváček, PhD.

Abstract

House price developments have a large impact on the macroeconomic stability, an example of which is the recent global crisis, partially triggered by a house price boom and bust. This work attempts to explain the behavior of house prices in ten Central- and Eastern-European countries over the last decade using three main methods: graphical comparison of the characteristics of the housing markets, panel data analysis, and time series analysis. First, a cross-country comparison shows that owner occupation rate or migration indeed play a role while other factors apparently do not. Second, the results of using the Pooled Mean Group estimator on a panel of all countries confirm that real income and unemployment are in general significant determinants of house prices. In the third part of the empirical analysis, VAR or VEC models are used on several individual countries to evaluate the role of national capitals as price leaders for the rest of the country. These models are finally employed again to test for significance of other fundamentals in several countries. The diversity of results leads to the conclusion that house price determinants differ widely across the analyzed countries, although this may be partially attributable to the unavoidable issues of variable house price data quality and availability, which limit comparability of the results.

JEL Classification: C32, C33, C52, G21, P22, R21, R31

Keywords: Central and Eastern Europe, house prices, housing markets, time series analysis, price leader effect

Abstrakt

Vývoj cen obytných nemovitostí má zásadní vliv na makroekonomickou stabilitu, jak dokazuje nedávná globální krize, při jejímž zrodu stálo prasknutí cenové bubliny na trhu nemovitostí. Tato práce si klade za cíl vysvětlit chování cen obytných nemovitostí v deseti zemích střední a východní Evropy v uplynulé dekádě. Hlavními použitými nástroji jsou grafické srovnání charakteristik trhů s nemovitostmi, analýza panelových dat, a analýza časových řad. Srovnání indikátorů trhu nemovitostí s vývojem jejich cen ukazuje, že procento lidí bydlících ve vlastním domě či migrace jsou s vývojem cen skutečně svázány, zatímco jiné faktory zřejmě nehrají podstatnou roli. Odhad metodou Pooled Mean Group na panelu všech zemí potvrzuje, že reálný důchod a nezaměstnanost jsou důležitými determinantami cen nemovitostí v tomto regionu. V další části empirické analýzy je pomocí modelů VAR či VEC zkoumána role hlavních měst jako cenových vůdců pro zbytek země. Tyto typy modelů jsou posléze použity ještě pro test signifikance dalších determinant v několika zemích. Rozmanitost výsledků vede k závěru, že se determinanty výrazně liší zemi od země, nicméně nelze vyloučit ani to, že zde hrají roli nesourodá kvalita a dostupnost dat o cenách obytných nemovitostí, které snižují porovnatelnost výsledků.

JEL klasifikace: C32, C33, C52, G21, P22, R21, R31

Klíčová slova: střední a východní Evropa, ceny obytných nemovitostí, trhy s obytnými nemovitostmi, analýza časových řad, efekt cenového vůdce

Contents

List of Figures

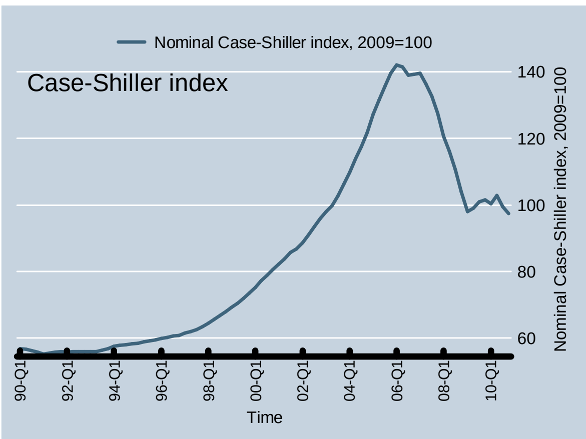
List of Tables

1 Introduction

The US subprime financial crisis marked the outbreak of the deepest recession since the Great Depression. In 2009, GDP of the European Union dropped by 4.3% and GDP of United States by 2.6%.[1] Even though the global crisis cannot be attributed to a single factor but – as pregnantly expressed by Bundesbank's President Weber (2008) – to a whole cocktail of causes connected with financial markets,[2] it would be impossible to explain the roots of the crisis without mentioning house prices.

According to Case-Shiller home price index[3] there had been a surge in house prices in the U.S. approximately from 2000 to 2007 - see Figure 1. Hand in hand with this growth went an increase in mortgages and other financial instruments, indirectly linked to house prices. For instance, banks used securitization transactions to eliminate credit portfolio risk from their balance sheets, thus decrease their regulatory capital requirements, be able to lend more and reach higher earnings. When the prices collapsed, many loans defaulted and the shock quickly spread to the holders of other instruments, linked to credit portfolios, mostly banks and insurance companies. Eventually, the situation became so grave that the US government had to intervene with a bail-out. (for a detailed study of the reasons and course of the crisis, see Breitenfellner & Wagner (2010) or Allen & Carletti (2010))

Figure 1: Nominal Case-Shiller index of home prices in the U.S.



Source: Standard&Poor's

[1] Source - World Bank Databank, http://data.worldbank.org/

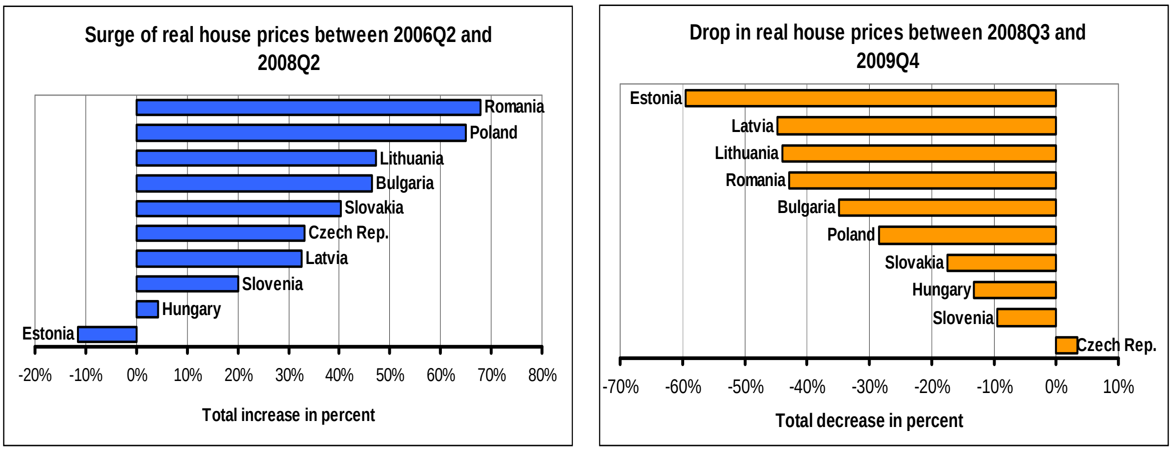
[2] Weber names 3 main reasons: lax lending standards, weaknesses in credit transfer, and overly optimistic assessment of structured securities

[3] A prominent price index for American residential housing market, compiled by Standard&Poors agency at 1-month frequency. According to the agency, it covers 75% of the whole housing stock by value.

This recent event illustrates that house price bubble bursts pose a large threat to the whole economy. Although the present author considers it highly improbable that a similar scenario could happen in Central or Eastern Europe, Figure 2 shows that there indeed has been a house price boom and bust in most of the countries. Even though the story of the U.S. can hardly repeat, if only because the mortgage market in CEE is much less developed, house price bubbles pose a problem to any market economy. If a bank overestimates the value of collaterals, its calculated Probability of Default (PD) and Loss Given Default (LGD) will probably be underestimated. If the bubble bursts and creditors default, the bank may have problems with liquidity or even become unable to meet its liabilities. If its creditors realize this, they may engage in a bank run. A bank run on a large scale can lead to the crash of the whole financial system. Therefore it is important for governments to identify periods when house prices are misaligned from their fundamentals so that they can include this consideration in the fiscal and monetary policy (prices usually surge when, ceteris paribus, interest rates are low, as shown by Ahearne *et al.* (2005)), regulations (e.g. rent control), and stress testing (Hlaváček & Komárek (2009)). The objective of this work is to find the most important fundamentals, approximate their impact on house prices, and find how fast prices tend to equilibrium after a deviation.

The work is structured as follows: Section 2 provides motivation and introduction to the topic, including the basic model of house prices. Further it discusses the key features of housing markets in general that set them apart from markets of other durable goods, analyzing issues connected with these features and solutions to them. Section 3 abandons the bird's-eye view and focuses on specific features of housing markets solely in Central and Eastern Europe, and their origin. Then we move even closer to differences among CEE countries. A graphical analysis is performed in order to find a connection between the varying characteristics of housing markets here and the patterns of house price development. Section 4 uses panel and time series analysis in order to find and evaluate the main determinants of house prices. The explanatory variables range from the conventional ones, such as real output and unemployment, to less used ones, e.g. construction price index or direct investment. An attempt to confirm the "price leader effect" of the national capitals for the rest fo the country is also made. Section 5 concludes.

Figure 2: Boom and bust in CEE



Source: national statistical offices, central banks, real estate agencies

2 Understanding house price developments

2.1 Reasons to study house prices

The term *housing* encompasses multiple activities in the economy, influenced by house price developments: construction, real-estate development, bank services and mortgage bank services, appraising, etc. These activities comprise a relatively **large part of GDP**; it is estimated that residential investment alone constitutes in average 6.5% of GDP in advanced countries. (IMF(2008b) in Hilbers *et al.* (2008)). Further, an increased activity in housing sector boosts aggregate demand. Tax revenues from land transfer are also not negligible. (Eurostat (2011))

Houses constitute a **significant part of wealth of households**. Therefore **changes in house prices influence private consumption** via the wealth channel. (Hlaváček & Komárek (2009)) This influence can be both **direct** and **indirect**.

In the first case, when house prices grow, a person can sell his apartment and buy a smaller one for a relatively lower price, or sometimes even an equally desirable apartment for a lower price. For example, in Singapore some people can *"sell their public housing units in the resale market and purchase new flats at subsidized prices using below-market interest rate loans from the state."* (Edelstein & Lum (2004)) The spared money would then increase their consumption.

The indirect effect on consumption manifests itself when a house is used as a collateral - a more valuable collateral means a higher purchasing power to finance consumption expenditures. This effect plays a larger and larger role as financial services connected with housing are becoming more widespread.

In spite of these facts, the overall effect of changing house prices on consumption is unclear. We can find counterarguments both in case of the direct and indirect effect. As for the direct effect: at one point in time, those who want to "trade down", i.e. move to a smaller house, will benefit from increasing house prices (if prices of all houses increased by 10% and I wanted to move to a cheaper house, the price increase would have made it relatively cheaper), but those who want to "trade up" will be correspondingly worse off. (Miles (1994)) Then, if everyone's propensity to consumption was the same, the impact of rising house prices on consumption should be zero! As for the indirect effect: expectations that the price increase is only temporary could cancel out the effect.[4]

These counterarguments are not only theoretical -Chen (2006) in a study of Sweden confirms that if a house price change is anticipated as transitory, it has no significant impact on consumption in the short run, which suggests that the indirect effect is indeed not significant, at least under these conditions. On the other hand, Chen also finds that in the long run, disposable income, aggregate consumption and housing wealth move together, which says nothing about the character of interaction between housing wealth and aggregate consumption, but confirms that there is a relationship. Other studies go further, analyzing the impact of house price on consumption quite successfully (e.g. Campbell & Cocco (2007)); therefore we cannot reject that this influence does exist.

House prices also have an **effect on well-being**. (Ratcliffe (2010)) Higher house prices usually point to higher housing quality. They are also correlated with better public services. On the other hand, growing house prices shift wealth from non-homeowners to homeowners. This should also shift general satisfaction from homeowners to non-homeowners. Surprisingly, Ratcliffe's results show that in periods of rising prices, both groups' well-being increased. The possible explanation could be that trends in house prices reflect the general prosperity of the economy,[5] which has a positive effect on everyone.

---

[4]This resembles the Barro-Ricardo equivalence, which describes a situation where a decrease in taxes will not boost consumption, as the consumers view it only as transitory, expecting that the government will have to increase the taxes again in the future to pay off the deficit, so they spare the extra money for these times. Similarly, house owners will likely not take a larger mortgage and increase consumption, if they expect the value of their property, used as a collateral, to sink: it would no longer cover their mortgage. For the same reason, the willingness of banks to provide mortgages would also be lower if they considered a price increase to be only transitory. In both cases, the stagnant volume of mortgages would have no effect on consumption.

[5]The general prosperity is well proxied by real GDP, which is considered an important determinant

2.2 Analyzing house prices

After explaining the importance of house prices we turn to their analysis. First we introduce a general model of house prices. As it would be impossible to analyze house prices without taking account of features of housing markets, we present the key features. Further suggestions for coping with issues connected with house prices are discussed. The arguably largest issue, availability of data on house prices, deserves its own part and concludes this Subsection.

2.2.1 General model

The housing market is usually modeled as two markets: one for the stock of existing houses, where the prices are determined, and the other one for the newly built houses, where the level of new investment is determined.(Cahlík *et al.* (2010)) The house price equation, associated with the first "market", is in most cases derived as an inverted demand equation. Here we present only a simplified derivation of a (also simplified) model according Muellbauer & Murphy (1997).

The housing demand for services of an individual or a household is represented by

$$\frac{H}{POP} = f(y, \mu, D), \tag{1}$$

where H is the demand for housing, POP is the number of individuals or households, y is average real income, μ is real user cost of housing services and D represents other factors that shift the demand curve. The real user cost can further be specified as

$$\mu = p_h(r + \delta - p_h^e/p_h), \tag{2}$$

where p_h is real house price, r is tax-adjusted interest rate, δ is depreciation rate or maintenance rate including property taxation, and p_h^e/p_h is the expected rate of appreciation of house prices. The user cost factors can be for simplicity written as one coefficient: $\upsilon = r + \delta - p_h^e/p_h$. Then $\mu = p_h\upsilon$. Plugging into 1 and inverting we get the equation for house price, determined on the market of existing houses.

$$p_h = g(\frac{H}{POP}, y, \mu, D) \tag{3}$$

of house prices.

This equation contains the most important factors that shape house prices. Some are already explicitly used (y or r), many others are "behind" the variables (demographic factors are included in POP, credit constraints are a part of user costs μ). The equation also gives us the house price, which then enters the market for the flow of construction (or more generally, market for housing services:

$\bar{H} = F(p_h, POP, S) - \delta H,$

where S represents other factors that shift the supply curve.[6] Because of the high inelasticity of housing supply it is mostly the demand side factors that are analyzed because they play a much more important role in determining house prices. Nevertheless, in the long term housing supply also plays a role: if it is too inelastic, it contributes to the creation of house price bubbles.

2.2.2 Specific features of housing markets

Houses are important commodities – they are often the most important component of their owner's wealth and the value of total residential buildings in most developed countries is several times that of aggregate annual output. They have a number of characteristics that distinguish them from other commodities (the list is based on Miles (1994)):

- *Durability* - A significant part of total housing stock in developed countries was built more than 100 years ago.

- *Uniqueness* or *Heterogeneity* – No two houses or dwellings are exactly the same. Although it is questionable whether there is any difference between two identical family houses in "Strawberry Street", as developers like to name new streets, or between two uniform apartments in a block of flats, each dwelling is unique at least in its position or address. These but also other differences (in size, equipment, etc.) can be rather small but also immense. The irreplaceableness of houses means they have to be treated differently than other tangible assets.

- *Inelasticity of supply* – This can be understood in two ways: First, it means that nobody can purchase a particular house if the owner is not willing to sell it, so the buyer has to settle for a different one. Second, it relates to the possibility and speed

[6] These include offering prices, investors' expectations, the stock of residential buildings or construction - the flow of new houses.

of construction of new houses. Supply can lag behind demand impulses because of the time necessary to find new land, obtain the permissions, find financing, and construct the new building. If demand sinks again during the construction period, the supply response will also be lagged. (Hilbers *et al.* (2008))

- *Collateral* - The possibility to raise loans against housing collateral (taking a mortgage) is greater than for most other assets; people can use their present house as collateral for a loan but they can also buy a new house and use it as collateral. These loans have in general more favorable conditions than loans for other purposes.

- *Developed secondary market* – A significant part of housing market consists of the existing stock. Turnover on this market is relatively low due to significant transaction costs, which are usually greater than 8% of the purchasing price.

- *Financial intermediaries* – primarily commercial banks, mortgage banks, housing finance specialists and others – are more involved in the housing market than in other consumer markets for durable goods. Conditions of sale differ widely because the contracts usually result from bilateral negotiations. These include agreements on the price, sale conditions (e.g. distribution of costs), or arrangements regarding the condition of the house (selling after a repair). (Hilbers *et al.* (2008))

- *Taxes* - Houses are also subject to differential tax treatment and subsidies: in dependence on the housing policy, the government can either support home ownership or renting.

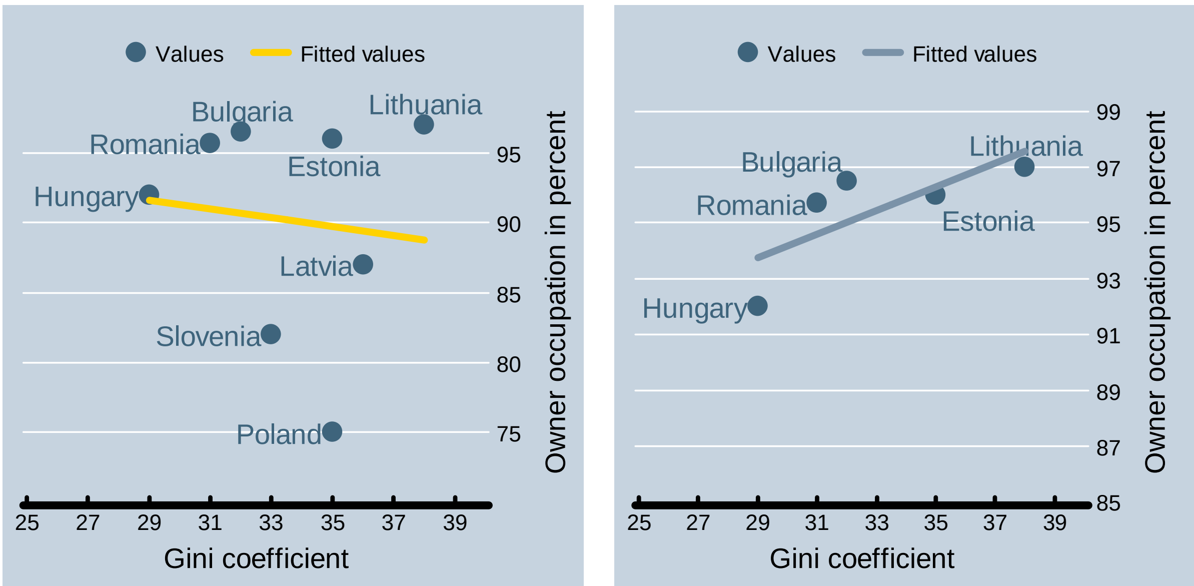
- *Price volatility* - And last but not least, although house prices are usually by far not as volatile as, for instance, stock prices or exchange rates, their implications for distribution of wealth makes even relatively mild changes relevant. The effect of increasing house prices can act in both ways – if the proportion of owner occupation[7] is high, the ownership of wealth will equalize, but in the opposite case, inequality can be expected to rise. (Miles (1994)) Yet can we observe this even in reality? We performed a simple graphical analysis, comparing owner occupation and Gini coefficient,[8] where available, in one graph. We payed special attention to the year of

[7] In other words, homeownership - the percentage of population, living in a residential property they own.

[8] A coefficient between 0 and 1 that measures equality of distribution of income or wealth in the society,

issue, always using data from the same year, or an estimate based on close values. In Figure 2 we see therefore data from different years: 2002 to 2008, but that should not cause any distortion. The results for all CEE countries with available data is that there is indeed an indirect proportion between owner occupation and distribution of income[9]. Nevertheless, if we omitted three variables, the result would look much different. Drawing any conclusions is therefore impossible here. Using a larger sample of countries and testing them sorted into groups (e.g. according to geographical location, population or other characteristics) would shed more light on the issue, yet this is not in scope of this work.

Figure 3: Link between Gini coefficient and owner occupation



Sources: World Bank, EMF, own adjustments

2.2.3 Coping with the specific features

It seems self-understood that when we want to study house price determinants, we will simply take data about house price development and then try to find factors that could have influenced it. But unfortunately, trends in house prices cannot be captured as easily as, say, trends in oil prices. The main reasons for this are the **low turnover and house heterogeneity**.

based on Lorenzo curve; 0 is total equality, 1 means that all the wealth is owned by one subject. In the world, the lowest Gini coefficient and thus the highest equality has Sweden, 0.23, and the highest Gini coefficient has Namibia, 0.70, although not all countries have been assessed - e.g. those missing in Figure 2.

[9]The coefficient of the regression line, even though insignificant, was -0.31; a country with owner occupation higher by one percent had on average lower Gini coefficient by 1/(0.31) = 3.22 points.

What house prices are we actually analyzing? The aggregate price of the housing stock, or only of its subset? We would like to know the former. *"The value of a house is most appropriately defined by the value it could fetch under current market conditions."* (Palacin & Shelburne (2005)) This value is usually expressed in price per square meter. To find the value of the housing stock, we would need to offer the whole housing stock and analyze the price buyers would be willing to offer. For both logical and technical reasons, this is not possible. The only prices we are able to observe are those of currently traded houses. What we can do is to record the price trend of recent transactions and try to translate it into the trend in value of the whole stock. But there are problems connected with this "translation".

Among the problems are the heterogeneity and immovability of the housing stock. If all houses were the same and tradable, arbitrage would push their prices roughly to the same level in the whole country. Nonetheless, no two houses are the same, so a perfect translation from the traded houses to all houses is impossible. What methods are used to overcome this limitation?

One option is the so-called **stratification** or in other words **"mix-adjustment"** (Eurostat (2011)) - dividing houses into groups according to certain characteristics (size, location) and calculating the value of the whole housing stock by multiplying the average recorded transaction price of houses in a group by the number of dwellings in the whole group. The price trend can then be calculated from the development of the weighted average of these prices. A limitation of this method is that it requires data about the character of housing market. There is also the question how precisely the groups ("strata") should be defined – if the groups are too large, compositional changes will affect the indices; if they are too small, there can be a higher sampling variability. (Eurostat (2011)) Nevertheless, this data is available only in few countries, which limits its use for comparison: e.g. the Czech Statistical office publishes quarterly price index of property separated by type, size of town, and condition.

Another way to estimate price trends is **observing repeat sales of only a small sample of houses**. This method is quite easy to implement but it only uses data on houses sold at least twice and the properties sold repeatedly may not be typical. (Palacin & Shelburne (2005))

In praxis, the published data are often only non-adjusted average transaction prices, which can suffer under composition bias. (Palacin & Shelburne (2005)) This means that some types of housing may be overrepresented in the sample in certain regions or time periods, which then distorts the real house price trend – e.g. if consumers move from buying cheap flats to family houses that will be more expensive, there may wrongly seem to arise a price bubble.

Apart from the problems connected with correct recognition of price trends there is the question of interpreting them. Do booming prices automatically mean a price bubble? Certainly not. The price increase can be an effect of increased quality of housing, as well as of a quicker turnover in the high-end market. (Hilbers *et al.* (2008)) The former will be analyzed in detail in Subsection 3.2.

2.2.4 Data issues

We described problems inherently connected with interpreting house price trends, independent on the size or character of housing market. Now we will focus on issues connected with the collection of data.

In most countries, the statistical offices or central banks collect some sort of data on house prices. The differences in *sources of data*, as well as *methodology* used, hinder direct international (or often even inter-regional) comparability.

There are several important *sources of data*, whose reliability varies widely. As mentioned in the previous part, one way to capture price developments is **following transactions**, i.e. recording the realized value. This price can be calculated from real estate tax returns that are connected with every market property transfer.

Another commonly used way of measuring house prices is **recording the asking prices**. This data is often published by real estate agencies. Although it may reflect price developments, its accuracy is doubtful – the asking price may be significantly different from the eventual transfer price. Moreover, relatively high asking prices can stay in the agency's catalogue for a long time, while cheaper houses are actually sold. Therefore it seems proper to possibly avoid indicators based on asking prices.

Other less frequently used methods include **hedonic regression**, based on evaluating houses by their attributes or characteristics, and **appraisal-based methods**. (for a complex description and evaluation of individual methods, see Eurostat (2011))

Methodology used in collecting and processing data cannot be ignored when the data should be compared. Yet the method of collecting the data is only one of the aspects necessary to have on mind. What area does the data cover? What currency is it recorded in, or is only a price index created? What kinds of dwellings does it cover? Are new houses calculated separately from the existing stock? How are reconstructions accounted for? Is the price per square meter, cubic meter, or the whole dwelling? Is the cost of installed equipment taken into account? Are the dwellings sorted? And if an aggregate index is calculated from the sorted data, how are the items weighted? These are only some of the many questions that need to be asked when comparing data from different sources.

From this example it should be clear that creating a standardized methodology for recording house prices would be very useful. This is not only the present author's opinion: Eurostat is currently developing a handbook on residential property price indexes (Eurostat (2011)) that should unify the methods of monitoring house prices and give guidance to those that have not yet been compiling a residential property price index. That would enable a qualified international comparison, which could be used for designing consistent policy measures on an international level, for instance. Another purpose of the unified index is the planned including expenditures on owner-occupied housing into the Harmonised Index of Consumer Prices (HICP).[10]

2.3 Related Literature

There is a wide range of literature related to determinants of house prices. Most of it is based on empirical findings. We can find a number of publications that focus on house prices in particular countries or groups of countries. Nevertheless, if we look solely for papers concerning house prices in Central and Eastern Europe (CEE), we will not find but a few exceptions from recent years.

We first introduce literature that is most closely linked to this work, i.e. papers studying determinants of house prices across a group of countries. Then we mention a few papers that analyze house price determinants in specific countries. Two works that describe housing markets and their specific features conclude this overview.

[10] The handbook is still under development (although it already has over 200 pages) and it remains uncertain whether it will be implemented, where, and how long it will take.

Most of the group studies we found focus on OECD countries. An indisputable advantage of choosing these countries is that there are much longer time series to analyze than in the case of CEE, for instance. Data issues remain an obstacle to those trying to analyze house price determinants in CEE countries - see Égert & Mihaljek (2007).

The kernel of empirical studies on this topic is usually as follows: choosing an econometric model and a set of explanatory variables, which are then used to explain the house price development in the chosen area. The explanatory variables are often called "fundamentals" as we consider them the "fundamental power" that indirectly moves the house prices.

The studies differ mostly in terms of size of the group of countries, methodology, and set of explanatory variables chosen. However, the main goal is similar: to identify the fundamentals that could explain the house price development. Success in reaching this objective naturally depends largely on which variables we choose - or are able to obtain. Real disposable income or real GDP, interest rates and unemployment belong to the most often used fundamentals. Depending on the models, data, variables, etc. various studies lead to diverse results for the elasticities of these "traditional fundamentals". In the study of 18 OECD countries Terrones & Otrok (2004) found that short-term real interest rate elasticity of real house prices varied from -0.5 to -1.0 while Sutton (2002) calculated it to be between –0.5 and –1.5 in a group of 6 OECD countries. Annett (2005) or Tsatsaronis & Zhu (2004) get the same sign yet the figures differ. The case of real disposable income elasticities of real house prices in all studies listed in Girouard *et al.* (2006) is similar – they all have a positive sign although the figures vary. That proves that there is probably no question about the effect of these fundamentals on house prices. How about other, more specific factors?

Stepanyan *et al.* (2010) show that sometimes we need to look beyond the most common fundamentals. In their study of 12 countries of former Soviet Union they used workers' remittances and foreign inflow as the explanatory variables of house prices. In contrast to what the results in OECD countries probably would be,[11] these factors proved sig-

[11] To the author's knowledge, there has been no study of OECD countries that would include remittances or foreign inflow among explanatory variables.

nificant. As another example, Égert & Mihaljek (2007) in their study of CEE countries chose the EBRD[12] transition indicators as proxy for the development of housing markets and housing finance institutions, and growth of real wages as proxy for improvements in housing quality. Both of these were significant. Other quite often used fundamentals include demographic factors (such as change in labor force or unemployment), the growth of housing credit, and equity prices.

As for methodological differences, it is necessary to point out that there are not only regressions but also other measures that can help us explain house price developments. The price-to-income and price-to-rent ratios are very commonly used as a measure of price sustainability.

Comparing developments in multiple countries usually includes one of types of regression. While some authors prefer ordinary least squares method (OLS) or its dynamic version (DOLS), others used a vector autoregressive model (VAR) – see Sutton (2002).

The choice of model is not arbitrary; each option has its advantages and drawbacks. For example the VAR model, as Sutton explains, allows for *"studying the dynamic influences of a small number of key determinants on home values"*, i.e. on house prices. It also enables us to compute the impact of unexpected shocks over a longer time period (several quarters in Sutton's paper). On the other hand, as Iacoviello (2000) cites from Cochrane (1994), VAR models cannot well distinguish between fundamental and non-fundamental variables, such as price bubbles. It also considers the response of house prices to changes in fundamentals to be linear while in reality it may be non-linear. An extension of the VAR model, which is also used in this work, is the vector error-correction model (VECM). In contrast to VAR, it allows to capture long-term equilibrium relationships between variables. (Cipra (2008))

Panel OLS models of different kinds are also employed by economists analyzing a group of countries. Égert & Mihaljek (2007) used panel DOLS model, Annett (2005) used fixed-effects (Least Square Dummy Variable – LSDV) OLS model. Authors of both of the studies found reasons for using their particular models – for a more detailed argu-

[12] European Bank for Reconstruction and Development

mentation, see the respective papers.

The last method we will mention is the Pooled Mean Group (PMG) estimator, used e.g. in Kholodilin *et al.* (2007) or Stepanyan *et al.* (2010). According to Pesaran (1997), PMG is a useful alternative between estimating separate regressions and conventional fixed effects estimators.

Now we will tackle a few papers that analyze house prices in only one chosen country. They can be used as an inspiration. Some parts of the analysis, such as region comparison, would be cumbersome to apply across a group of countries, but others may not.

Pagés & Maza (2003) made an attempt to find the dynamic relationship between house prices in Spain and a number of fundamentals by applying the VECM. They strived for choosing the most suitable data: for instance, as income they used household gross disposable income per inhabitant older than 25 years. They did this because the number of people over 25 per household has not almost changed over the studied period (1978-2002), while the total count of inhabitants per househould had been steadily decreasing.

Hlaváček & Komárek (2009) studied housing price bubbles in the Czech Republic. Besides the traditional explanatory variables they used quite uncommon ones, such as the rate of marriages and rate of divorces. The logic behind this was that both with a marriage or a divorce arises a demand for a new home, pushing house prices up. Divorces proved significant in several cases, although this result has to be taken with caution because of the relatively short time series. Another interesting thing in the paper was the attempt to recognize the capital, Prague, as a "price leader" of the CR. Using time series regression analysis, the hypothesis was not rejected – 1 or 2 quarters' lag of prices in CR compared with Prague was significant in most cases. We take inspiration of this analysis; the results can be found in Subsection 4.3.

The housing market has its specific features that need to be taken into account. Miles (1994) enumerates 9 of them (see Subsubsection 2.2.2). He also provides a deep look into housing in context of the wider economy, discusses econometric issues, and uses empirical evidence to illuminate some questions. Palacin & Shelburne (2005), on the other hand, focus solely on the characteristics of private housing market in Central and

Eastern Europe, Armenia, Kazakhstan, and the Russian Federation. In their study the authors do not use econometric models but a rather simplistic description of the current state of private housing market and its history. They also compare the markets with the developed and try to understand why some differences persist.

House prices are a relevant topic in 21st century, as the count of new studies shows and as the world crisis proved. The amount of literature corresponds to this fact. Yet from this overview it is clear that there are regions which had long been negliged, such as CEE. This work tries to help fill this gap.

3 Characteristics of CEE housing market

Up to this point, the discussion about assessing house price trends, data collection, and methodological issues has been mostly general. Now we will describe problems connected with house price data specifically in CEE.

3.1 House prices in CEE: factors specific to house price development

The housing market in CEE differs from the one in most OECD countries. Its today's characteristics are highly influenced by the socialist history of these countries, the transition and related institutional development. Nevertheless, it is necessary to note that not all the characteristics can be attributed to the socialist past – e.g. owner occupation varied widely across CEE even before transition. The characteristics have an impact both on the absolute level of house prices and their dynamics so we cannot omit them when looking for house price determinants – they can be equally relevant for understanding price developments as the conventional fundamentals. First I describe the roots of the specific features and then their impact on current house price development.

3.1.1 Historical background

Before transition:

- There was **no housing market** in the current sense but instead, **three basic patterns of housing ownership co-existed**: state-owned, cooperative, and privately owned. Yet in private housing, there were regulations regarding the number of dwellings a family could own and the sale price so the ownership was not really private in today's sense. The owners of the other types also had limited rights that are nowadays associated with ownership. Purchasing houses had to be approved by government but after that housing loans were provided automatically after fulfilling given conditions. During transition, a large part of the housing stock was privatized – partially by restitution, partially by selling it to other individuals, usually the present tenants.(Palacin & Shelburne (2005)) The selling conditions were very favorable so in the early 1990s, the share of owner-occupation was very high in most CEE countries while the household debt was very low.

- The overwhelming **majority of construction was carried out by state**. The general quality of new housing was low: a large number of *"standardized, low-quality, prefabricated apartments in pre-cast concrete high-rise buildings in high density zones on the periphery of cities"* (Palacin & Shelburne (2005)) was built. In spite of this construction focused on quantity rather than quality, the housing stock was low relative to population; there were chronic shortages of housing with long waiting lists for apartments. The legacy of limited and low-quality housing helps understand today's trends in the current housing market.

During the transition:

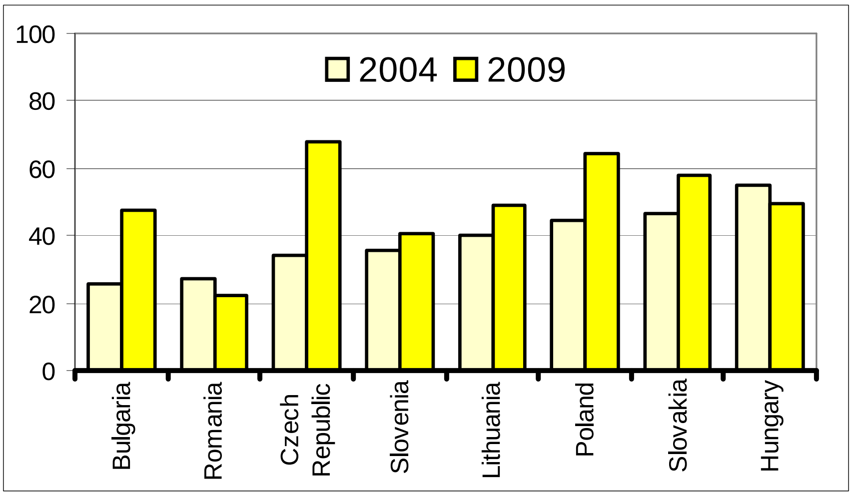
- House prices rose swiftly, partially because of surging inflation. Nevertheless, there is evidence that they grew slower than inflation. Price-to-income ratios dropped between 1990 and 1994. Residential construction collapsed during 1990s because state stopped supplying it, the private housing finance was yet undeveloped, there was a high level of uncertainty due to insufficiently specified property rights, and the economic downturn connected with transition stifled effective demand. It often took more than a decade for residential construction to return to pre-transition levels (Palacin & Shelburne (2005)).

3.1.2 Impact of CEE's past on the current housing market

Experience from Western Europe around 1990 shows that house prices tend to rise towards equilibrium levels[13] when markets are deregulated. (Égert & Mihaljek (2007)) The housing

[13]By equilibrium we mean a situation when there is no systematic pressure on house prices, quantity demanded equals quantity offered (by home-sellers).

Figure 4: Proportion of residential mortgage debt to total household loans as per cent



Sources: Eurostat, EMF, author's calculations

restrictions in CEE were abandoned after the fall of socialism but this deregulation itself did not suffice for the prices to reach the equilibrium because the availability of housing credit was low, limiting the demand. Because of the sophisticated character of housing institutions, their development lagged behind financing institutions. The "drive towards equilibrium" is therefore probably still underway.

The price-to-income ratio had been low in 1990s in CEE partially because housing finance was undeveloped. Since then, the volume of mortgages – the most significant part of housing finance – has been growing in most countries of CEE (especially Estonia and Latvia) together with total households' credit; in most cases, it was actually a key part of total credit growth - see Figure 4 with countries where data is available. Housing financing belongs to important drivers of housing demand, therefore its development tends to push house prices up. As there is still a wide gap between volume of housing loans in CEE and in Western Europe, house prices can be expected to grow further as the relative volume of loans becomes comparable to the West.

For the residential construction investment it took quite long to pick up; the relative size of the housing stock remains low compared to Western Europe. That implies that the market is not yet saturated and when the number of dwellings gets to levels similar to the West, prices should decrease.

Small, poor quality dwellings, many of which were built in socialism, still dominate the current housing stock. (Palacin & Shelburne (2005)) In the past, homogeneity of construction unabled people to grow up the "property ladder" (move to a better dwelling);

this trend is changing as higher quality housing of various types becomes available. The relative prices of poor quality dwellings should sink as people choose alternative housing options. That would cause a larger price gap between little marketable, low quality housing, and the higher quality housing. This process could however be partially reversed by improving the poor quality housing by reconstructions.

The economic transition changed the geographical distribution of economic activity and population. (Palacin & Shelburne (2005)) Business thrived mostly in cities while in smaller towns and villages it stagnated. That had obvious implications for house prices. The trend of weaker house prices growth in smaller towns continues up to this day. As houses form a large part of households' wealth, the boom shifted wealth from owners of houses in smaller towns to owners in larger ones. The boom in larger towns also spurred working opportunities, enticing workers from other places; their demand for housing further pulled house prices in cities up.

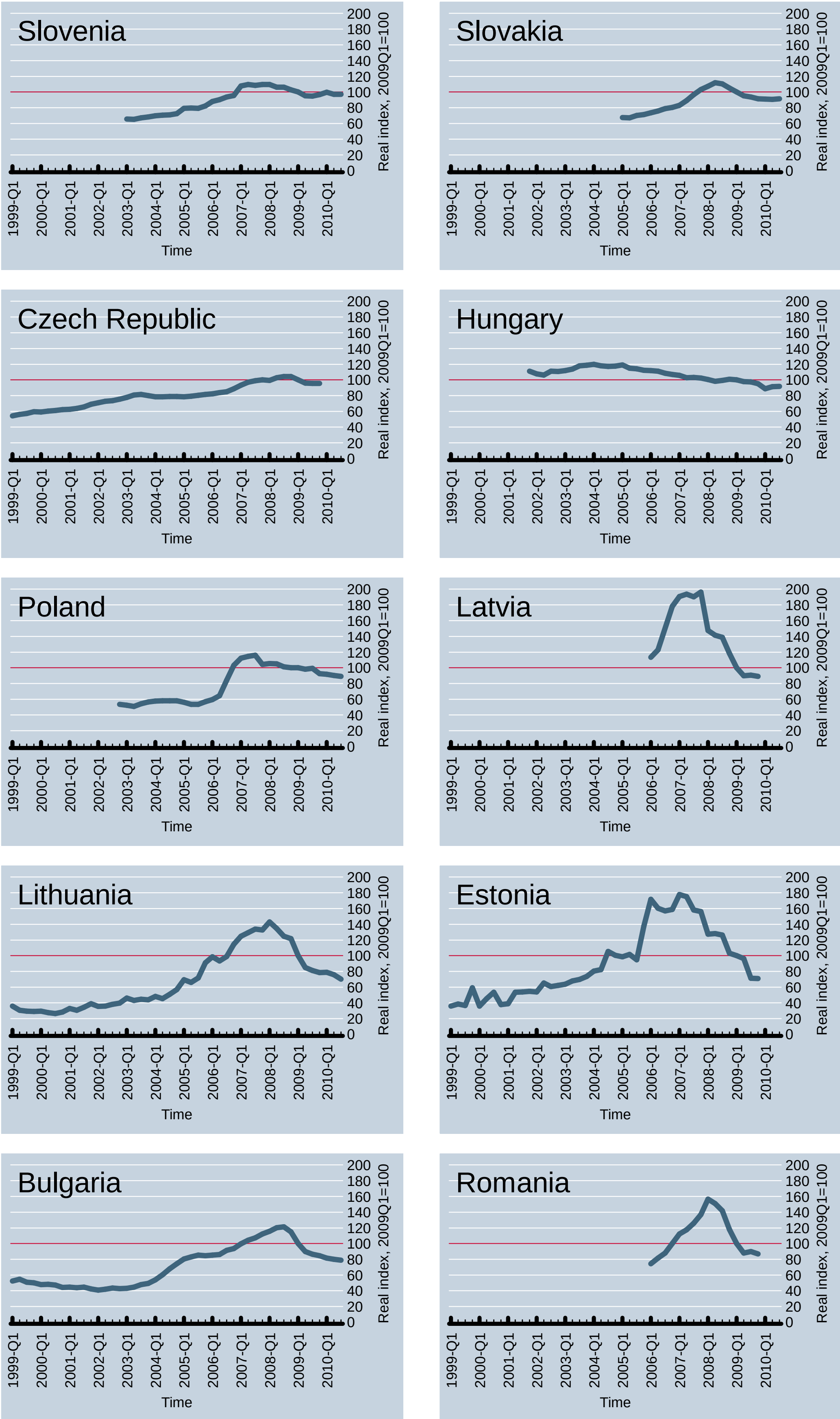
The high increase in house prices over the last two decades could have also been caused by the initial undershooting. During socialism, there was no market mechanism adjusting prices and house prices were very low compared to other durables, such as cars. Nowadays, houses are several times more expensive than middle-class cars.[14] This change would be consistent with the hypothesis that house prices were originally under the equilibrium and that the house price growth was only a correction independent on fundamentals. Nevertheless, we assume that this correction took place within 5, maybe 10 years since the fall of the Iron Curtain.

3.2 Recent developments

Here we describe the present state of the housing market and changes over the last years (earlier data are in many cases not available) in order to find any link between the development of house prices and characteristics of the individual housing markets. We start with the object of this study, house prices. From Figure 5 it seems that countries could be roughly divided into 2 groups: countries that have undergone a house price boom&bust and those that have not (Hungary, Czech Republic, Slovenia). In Section 4 we will try to

[14] Égert & Mihaljek (2007) note that at the beginning of transition, the relative price of a middle-class passenger car to a typical apartment in a block of flats in CEE was 1:1; this ratio reached about 1:4 by 2006.

Figure 5: Real house price developments in CEE

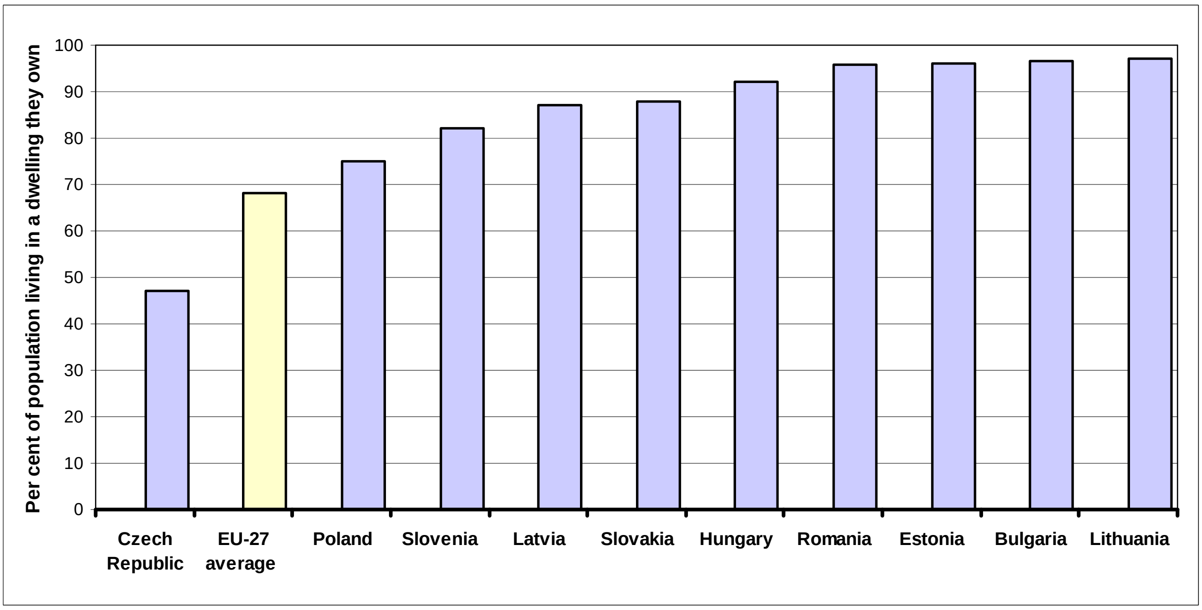


Sources: National statistical offices, Central banks, BIS, own calculations analyze whether this development was in compliance with fundamentals.

Among important characteristics of housing markets are the share of owner occupation, relative volume of mortgages and of total household debt, number of people per dwelling, quality of housing and inelasticity of supply interest rates on housing, credit rationing and others.

Owner occupation rate is high in all analyzed countries in comparison with Western Europe; the only exception is the Czech Republic. While it cannot be said from Figure (6) that there is a link between owner occupation rate and boom&bust development of house prices, it is worth noting that the country with the lowest ownership rate, Czech Republic, is the only country where prices had been growing even after 2008. A higher long-term demand for owner-occupation could be a reason for that.

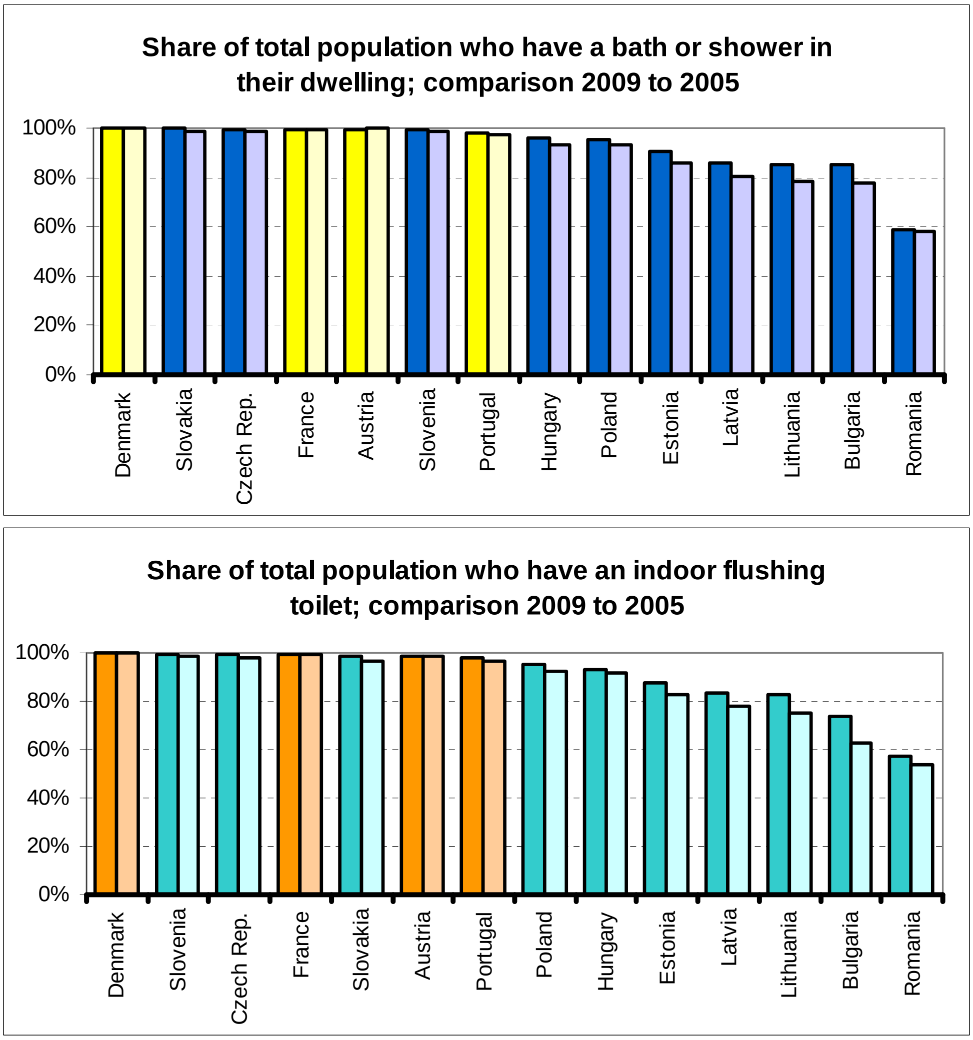
Figure 6: Owner occupation rate



Source: EMF

Quality of housing is uneasy to assess without specific data. Égert & Mihaljek (2007) compensate for this insufficiency by using real wages as a broad proxy. Here, however, we use a more unambiguous measure of housing quality: share of households who have a shower or a bath, and share of households with a flushing toilet. Graphs in Figure 7 show that these conditions vary widely across CEE. The measure is probably of little use in assessing development in more developed countries, where the proportion of equipped households reaches nearly 100%, but it certainly shows the improvement in housing quality in the rest of countries. A higher average quality of housing can partially explain the increase in house prices. It also moves higher the equilibrium price of dwellings and can be one of the reasons why the prices did not return to the pre-boom level but remained above it.

Figure 7: Indicators of housing quality

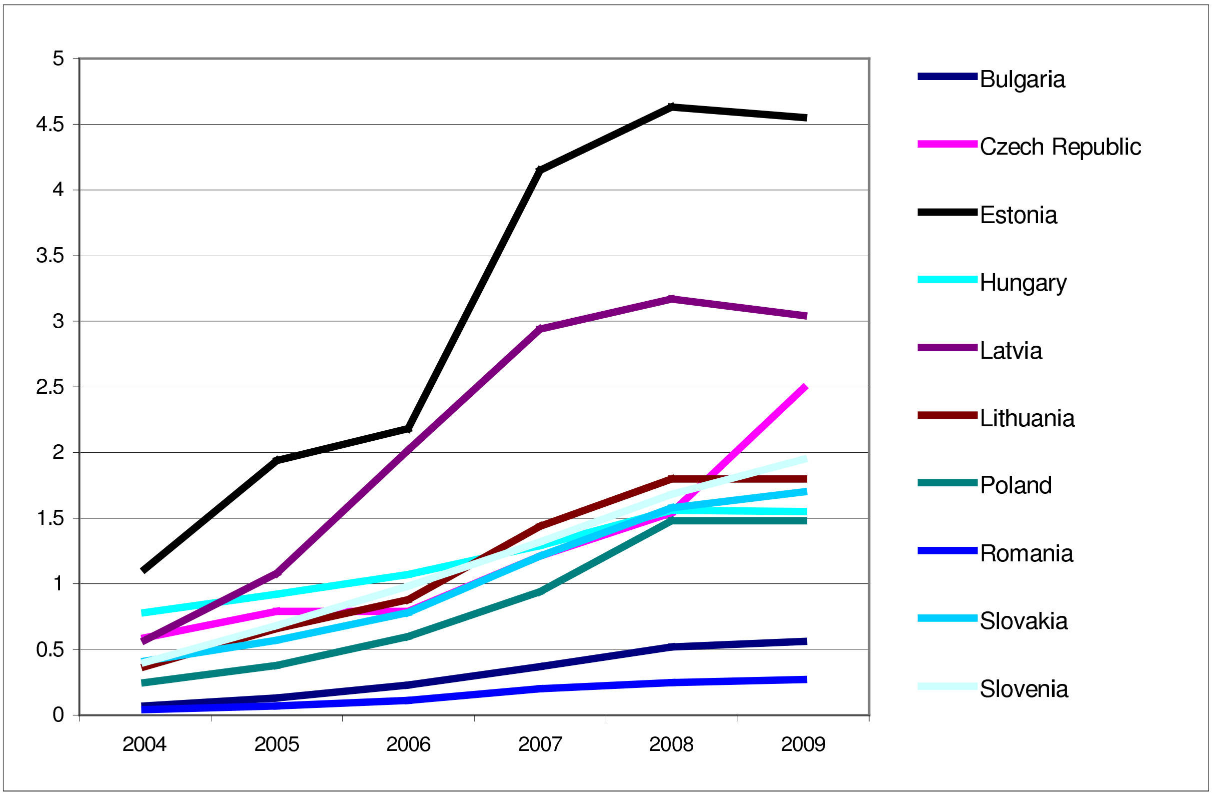


Source: Eurostat

Even in the more developed countries, most people cannot purchase a house only from their own savings. The availability of **housing loans**, in particular **mortgages**, is then an important determinant of the effective demand of households. As can be seen in Figure 8, the volume of mortgages has been growing quite fast in most countries; the least developed housing markets in this aspect are in Bulgaria and Romania.

The total volume of household credit had also been growing in this period. Figure 4 shows that mortgages played an important role in this growth – their share in total loans

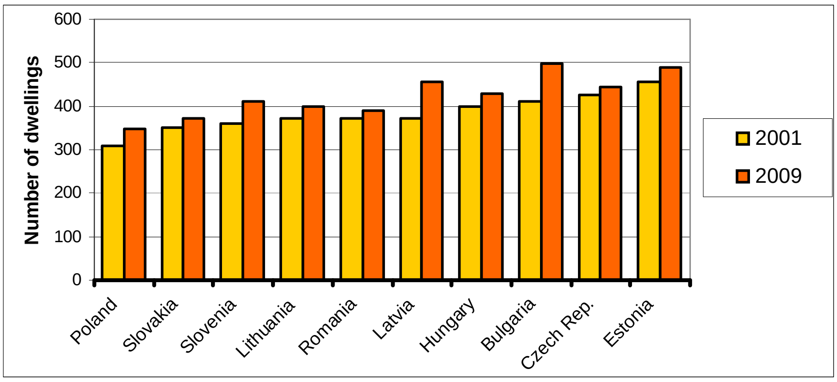
Figure 8: Residential mortgage debt per capita, EUR thousand



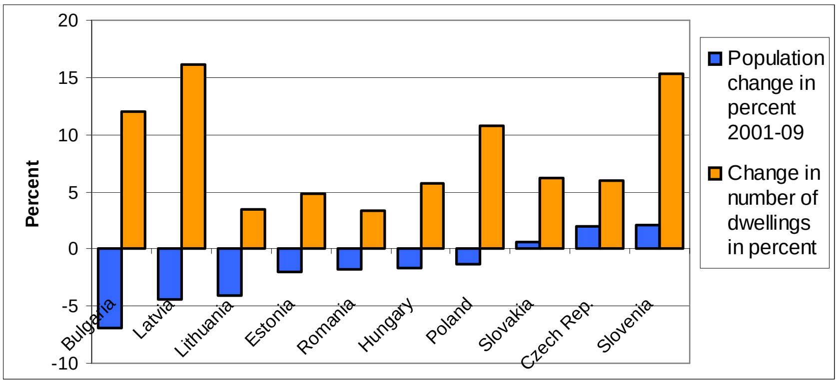
Source: EMF

increased in all countries besides Hungary and Romania. In Bulgaria, the high increase of mortgage share was largely due to a low level of total credit over the whole period. Further it is noteworthy that in all countries besides Hungary, Romania and Slovenia the mortgage growth accounted for more than a half of total household credit increase. Nonetheless, in spite of the speedy growth, the proportion of residential mortgage debt to GDP is still significantly lower in the sample than in most western countries; even the fastest growers of CEE, Estonia and Latvia, are still under the EU27 average of 51.9% of GDP. This indicates the potential for the housing finance market to grow further and push house prices up.

Figure 9: Housing developments between 2001 and 2009



(a) Average number of dwellings per 1000 inhabitants



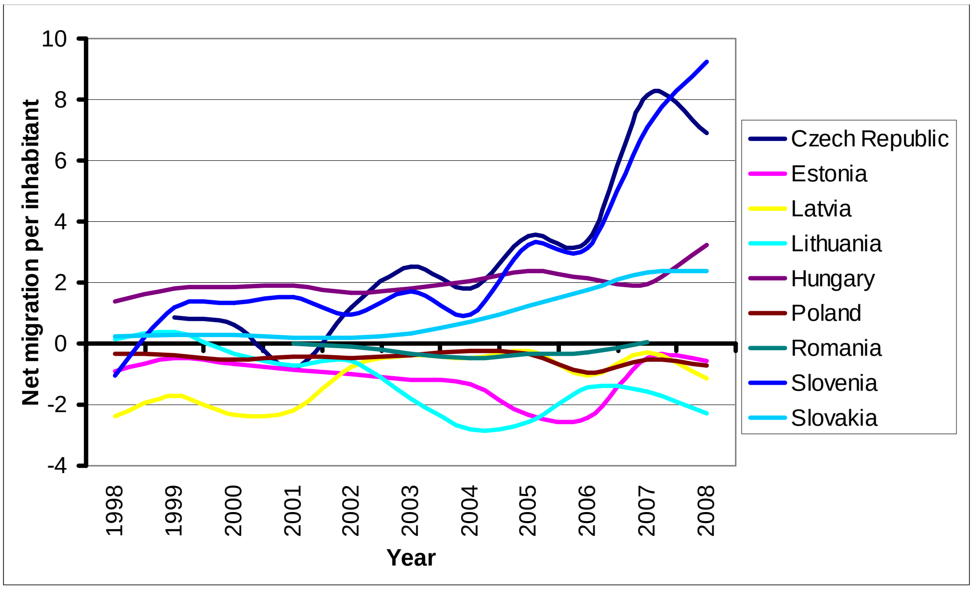
(b) Population change and construction

Sources: Eurostat, EMF, author's calculations

As already mentioned, an important characteristic of the housing market is the **inelasticity of supply**, which contributes to the volatility of house prices, keeping the market out of equilibrium for a longer time period. The reaction of construction to price developments will be analyzed later; here we only try to find any connection **a)** between the total size of housing stock relative to population and house price development, **b)** between population and construction, and **c)** between change in population and house price development.

a) From Figure 9a) it is hard to find any correspondence between the housing stock in 2001 and the subsequent development: the "boom&bust" countries are on both ends of the graph. As for question **b)**, there should be a correlation between change in population - which partially determines demand for housing - and construction, responding to this demand. In contrast to this logic, however, Figure 9b) shows that the connection can be spotted in only a few cases; some of the largest construction booms took place in countries where population sank significantly, namely Bulgaria and Latvia. The construction boom

Figure 10: Net migration per 1000 inhabitants



Source: Eurostat, author's calculations

can hardly be attributed to previously low number of dwellings per capita: these countries are in the upper half of this table.

The answer to question **c)** seems quite clear - in countries where the population grew, prices were more stable and nearly did not sink during the crisis. This could be linked with the general sentiment in the society, keeping house-sellers believing that the value of their house would not sink because of a lower demand caused by a lower population. But the effect may have been also in the opposite direction - a higher standard of living, which is connected with higher house prices, attracts immigrants. In Figure 10 you can notice that all the "no boom&bust" countries have had the highest net migration.[15] Whichever direction the causality is, the link between the Figure and house price developments gives us a clear incentive to perform an econometric test for a relationship, which we will do in Subsection 4.4.

This evidence altogether proves that if changes in population have an impact on house prices, it is through demand, not supply channel.

4 Empirical analysis

This section consists of 3 analyses: a) we try to find the main fundamentals behind house price dynamics first in the panel of all 10 countries, using the PMG model; b) we analyze

[15]By net migration we mean immigration minus emigration.

the role of national capitals as "price leaders", using the VAR model. c) we use a wider set of fundamentals, including volume of mortgages, in several countries separately, using the VAR and VECM model.

We obtained the house price data from national statistical offices, central banks, and real-estate agencies. An overview can be found in Table A1. All the data is subject to the methodological limitations, outlined in Subsubsection 2.2.4.

It is crucial to note here that analysis based on fundamentals assumes that relations between the fundamentals and house prices are stable. Nevertheless, in times of crisis it often happens that the economy moves "of its own accord"; models stop working, the ties between macroeconomic quantities are broken. Is it then possible to make an applicable model, if we base it on data from "normal times", as well as the times of crisis? And is it even desirable to attempt to create such a model, if it should only work in the "normal times" - when it is not much needed - but fails to work in case of unprecedented events?[16]

Some economists blame modern academic economics for *"constructing models that, by design, disregard the key elements driving outcomes in real-world markets."* (Colander *et al.* (2009))[17] There is a chance that the models in this analysis would belong to those, criticized in this paper. Yet as we can hardly model the *key elements*, by which the present author understands expectations, institutional setting, or investment sentiments, it would make more sense to simply search for house price determinants in other times than crisis. That could be reached by omitting the observations from the period of "boom and bust". Yet we have a few reasons that make this option unviable: first, dropping observations always deprives us of some information that could be important; second, our available time series are often too short; omitting observations of 2 or 3 years would mean dropping half of the data set. A time series analysis would then be impossible.

Lastly, it is important to note that even if the results exactly and correctly quantified the relationship between house prices and the fundamentals, it is mostly the policy maker who is in charge (apart from exogenous factors, such as the world crisis or investment sentiments). It is likely that the problem of house price bubbles begins with "GDP growth

[16]An ideal model would be one that would be able to predict these "unprecedented events". Unfortunately, its construction seems impossible; the best way to predict them is probably learning from the past and not repeating mistakes. In the context of the recent world financial crisis, that would include measures for a higher bank liquidity, using more transparent financial instruments and lending money prudently.

[17]The paper bears the incisive title *The Financial Crisis and the Systemic Failure of Academic Economics*, which speaks for itself.

bubbles", caused by immoderate government spending. A way to prevent property price bubbles should therefore begin with a more prudent macroeconomic policy. The analysis of house price determinants gives a clue as to how high the prices should be based on the fundamentals, but it cannot account for unmaintainable GDP (and other) booms. Therefore we could find ourselves in a situation where prices seem to be in equilibrium (based on the fundamentals) but as the fundamentals are themselves not in equilibrium, there is still a potential of a sudden bust, beginning for instance on the financial market and ending on the housing market.

4.1 Definitions

When analyzing time series, there are a few possible issues that need to be ruled out: stationarity and cointegration. First, however, we will shortly describe these issues and mention problems connected with them.

STATIONARITY

Definition 1

A stochastic process y_t is **covariance stationary** (or **weakly stationary**), if it satisfies the following requirements:

1. $E(y_t) = \mu,\ \mu < \infty$
2. $var(y_t) = \sigma^2,\ \sigma^2 < \infty$
3. $cov(y_t, y_{t-k}) = \gamma_k,\ k \in \mathbb{Z}$

These conditions can be summed up this way: a stochastic process is weakly stationary, if its mean and variance do not vary over time and are finite, and if the autocovariance (covariance between values of one series) between two values only depends on their distance from each other (k), but not on their location of the initial time period (t).

Definition 2

An **autoregressive process of order** p ($AR(p)$) is:

$$y_t = \alpha + \lambda_1 y_{t-1} + \ldots + \lambda_p y_{t-p} + \varepsilon_t,\ t = 1, \ldots, T$$

The disturbance term ε_t is assumed to be an independent and identically distributed (i.i.d.) random variable not correlated with y_t. Further $E(\varepsilon_t) = 0$, $var(\varepsilon_t) = \sigma^2$.

Obviously, the present value of autoregressive processes depends partially on its p past values. It can be shown that an autoregressive process is stationary, if the roots of its characteristic equation have modulus greater than one. Vice versa, if the modulus of any roots is less than one, the autoregressive process in non-stationary, which means that its statistical properties change over time; the time series is unstable. This means that we allow the relationship between the regressors and the regressand to change over time. But in such case, we would be unable to track any long-term relationship between them so the results of a regression would be of little or no use. Therefore we always search for stationary time series. Unfortunately, in reality many time series show a trend (income, consumption, inflation,...) and therefore are non-stationary. Nevertheless, some non-stationary time series can be transformed into stationary series.

Definition 3

A time series is said to be **integrated of order 1** ($I(1)$), if number 1 is a root of its characteristic equation.

An example of an $I(1)$ process is the **random walk with drift** (Greene (2003)):

$y_t = \mu + y_{t-1} + \varepsilon_t,\ t = 1, \ldots, T$

where μ is the drift. This can be rewritten as $y_t = \sum_{i=0}^{\infty}(\mu + \varepsilon_{t-i})$. If we take ε_t with a zero-mean and constant variance, we see that $var(y_t) = \infty$, therefore y_t is a non-stationary process. Nonetheless, the first difference already will be stationary:

$\Delta y_t = y_t - y_{t-1} = \mu + \varepsilon_t$

We can see that if a time series is $I(1)$, it is non-stationary, but we can make it stationary by using its first difference. There are a number of tests for unit root, which help to find whether a time series needs to be differenced or whether it is already stationary, e.g. the Dickey-Fuller (DF) test, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, the Phillips-Perron (PP) test, or the DF-GLS test, which is a modified Dickey-Fuller t-test in which the series has been transformed by a generalized least-squares regression.[18] Hadri's Lagrange Multiplier test is used for searching unit roots in panel data. We applied several of these tests with various results.

[18] See help in the program Stata.

COINTEGRATION

Definition 4

> Time series $a_t, \ldots, n_t$, integrated of order 1 are **cointegrated**, if there exists a non-zero linear combination of $a_t, \ldots, n_t$ that is stationary.

In other words, they share a common stochastic drift. Let us use an example with $i = 2$. If both y_t and x_t are $I(1)$, and $\exists \beta \, (p \times 1)$: $y_t - \beta x_t = \varepsilon_t$, where ε_t is white noise and is therefore $I(0)$, then they are cointegrated and the vector $(1, -\beta)$ (or any multiple of it) is its cointegrating vector. In this case, differencing the variables could easily prove to be counterproductive.

Cointegration enables us to distinguish between short-term (short-run dynamics) and long-term (long-run trend) relationship between y_t and x_t. After differencing the variables, we would be able to capture the short-term relationship but the problem here is that the long-term relationship would be obscured.(Greene (2003)) The tests used later in the analysis show that the regressed time series indeed are cointegrated, therefore we avoid differencing the time series and instead use econometric methods that deal with cointegration and differencing.

4.2 Panel data analysis

An important advantage of panel data over time series or cross-section data is that it allows controlling for individual heterogeneity. In this case, it means that it can "cope with" country-specific variables that often cannot be measured but which, if omitted, would lead to bias in the estimates. Another reason for using panel data is that decreases the problem of multicollinearity, common in time series analyses. It also "leaves" more degrees of freedom and reaches a higher efficiency. (for more advantages, see Baltagi (2005))

4.2.1 Estimation

We try to assess the role of several determinants of price p_h from (3) using panel regression. Because of the unavailability of data, the spectrum of variables that could be used for the whole set of countries is very limited. We considered the option of employing the model again with a richer set of explanatory variables only on sets of the countries for which both the house price data series and regressor series are available, but the possible groups would be unviable; a panel with BG, CZ, EE, and LT would be highly heterogeneous. Moreover, the PMG model is designed for models with a larger n. (which is here the number of countries)

As regressand we use real house prices, calculated from nominal house prices (or in three cases, from nominal house price indices) by changing to euro and then deflating by HICP of the euro zone. This operation was performed to obtain comparable data. Further we used two indices with 2009Q1=100 as a base: real GDP (volumes) and HICP of individual countries, and unemployment in percent. Real GDP is strongly correlated with real income, which increases consumption. Rising price level naturally increases nominal house prices, but Poterba (1984) argues that it also increases real house prices because of it decreases user cost via tax subsidy, such as exemption of housing capital gains for the elderly. Unemployment reflects the aggregate demand in the economy and influences disposable income. The equation can be simply written in this form:

$$p_h = g(\overset{+}{RGDP}, \overset{+}{HICP}, \overset{-}{UNEMP})$$

4.2.2 Model specifications

There are several basic versions of panel data models, which vary mainly in their assumptions regarding the cross-sectional and time component, and whether the intercept is correlated with the regressors. (see Hlaváček & Komárek (2009))

We employ the pooled mean group (PMG) model by Pesaran (1997). The PMG model was used in other studies of house prices, Kholodilin *et al.* (2007) and Stepanyan *et al.* (2010). There are several reasons why this model fits our purposes: it is well suited for heterogeneous panels with a relatively large number of groups and observations (in this case, $n = 10$ is not large but also not too low; T varies from 16 to 47, with an average of 39 per country, which is relatively large), which should be of the same order of magnitude, which is roughly fulfilled.[19]

Another advantage of the PMG model is that it provides a compromise *"between fixed or random effects model that requires all slopes to be identical across groups and the very general model where the slopes are treated as completely unrelated."* (Pesaran (1997)) The PMG model, in contrast, allows intercepts, short-term coefficients and error variances

[19] We understand the order of magnitude as the integer part of a common logarithm of the number; e.g. the order of magnitude of 47 is 1 because $log(47) = 1.67$; the order of magnitude of 16 is also 1 because $log(16) = 1.2$.

to differ across groups, while it constrains long-run coefficients to be the same. These restraints are probably closer to reality – they allow for an individual speed of return of groups to the steady state, but assume that in the long-run relationships between variables across countries are stable.

We assume the long-term relationship between house prices and fundamentals to be:

$$HP_{it} = \theta_0 + \theta_1 RGDP_{it} + \theta_2 HICP_{it} + \theta_3 UNEMP_{it} + \mu_i + \varepsilon_{it}, \tag{4}$$

where $i = 1, ..., 10$ and $t = 1, ..., T$ indicate country and time; HP is natural logarithm of real house price index, $RGDP$ is natural logarithm of real GDP index, $UNEMP$ is rate of unemployment and μ_i is country specific fixed effect. As already mentioned, the first two variables are indexed using 2009Q1=100. If the variables are I(1) and cointegrated, then the error term ε_i is I(0) $\forall i$. (Pesaran (1997)) These two issues are discussed in the next section.

The PMG model is based on the *autoregressive distributed lag* (ARDL) model, which provides a useful compromise between finite lag models and the geometric lag model.(Greene (2003)) The ARDL models use both the lagged values of the explanatory variables x_t (the *distributed lag* part) and lagged values of the dependent variable y_t (hence *autoregressive*) to explain the dynamics of y_t. The general form of an ARDL(p,q) model is

$y_t = \alpha + \sum_{i=1}^{p} \lambda_i y_{t-i} + \sum_{j=0}^{q} \beta_j x_{t-j} + \varepsilon_t,$

where $t = 1, \ldots, T$ is the number of time periods, y_t $(T \times 1)$ is a vector of the dependent variable, x_t $(T \times 1)$ is an explanatory variable, α is the intercept, coefficients $\lambda_i$and β_j are scalars, ε_t $(T \times 1)$ is the vector of disturbances. There can naturally be more than one explanatory variable; then we write ARDL$(p, q, r, ...)$, where p is the number of lags of y_t (which should be based on how old values still have an impact on the present) and $q, r, ...$ are numbers of lags of the respective regressors. The respective variables $(p, q, r, ...)$ are naturally independent of each other but in order to keep the calculations simple, we take the case where the number of lags of the explanatory variables is the same, i.e. ARDL$(p, q, ..., q)$. We also omit the intercept. Further, the PMG model is designed for panel data, therefore the general form will be even more different from the original ARDL(p,q) model:

$$y_{it} = \sum_{j=1}^{p}\lambda_{ij} y_{i,t-j} + \sum_{j=0}^{q}\beta_{ij}^T x_{i,t-j} + \delta_i^T d_i + \varepsilon_{it} \tag{5}$$

where $i = 1, \ldots, N$ is the number of groups (in this case, $N = 10$ countries), $t = 1, \ldots, T$ are time periods, λ_{ij} are scalars, β_{ij} $(k \times 1)$ and δ_i $(s \times 1)$ are vectors of unknown parameters. Further, x_{ij} $(k \times 1)$ and d_i $(s \times 1)$ are vectors of explanatory variables, where x_{ij} vary both over time periods and groups, whereas d_i only vary over groups.

In our case, there is no explanatory variable corresponding to d_i. Instead, we added the coefficient μ_i, which allows for the country-specific fixed effect. We assume that the impact of GDP, inflation and unemployment on house prices only demonstrates itself immediately or with a lag of one quarter. Further tentative regressions supported this assumption, with zero- or one-quarter lag significant but two-quarter and further lags insignificant.

By filling in our variables in (5) and setting $p = q = 1$, we get the ARDL(1,1,1,1) dynamic panel specification of (4):

$$\begin{aligned} HP_{it} = \lambda_i HP_{i,t-1} + \beta_{10i} RGDP_{it} + \beta_{11i} RGDP_{i,t-1} + \beta_{20i} HICP_{it} + \\ + \beta_{21i} HICP_{i,t-1} + \beta_{30i} UNEMP_{it} + \beta_{31i} UNEMP_{i,t-1} + \mu_i + \varepsilon_{it} \end{aligned} \tag{6}$$

The indexes of coefficients β indicate the number of the regressor, the lag, and the country: $\beta = \beta_{kji}$. Esthetically this seems more convenient than to write it as in 5 and add k to mark the regressor, i.e. β_{ijk}, for instanceβ_{i02}.

It is convenient to re-parameterize (6) in the following way: (for simplicity, we omit $HICP$ and $UNEMP$ here)

$$HP_{it} = f\lambda_i HP_{i,t-1} + \beta_{10i} RGDP_{it} + \beta_{11i} RGDP_{i,t-1} + \mu_i + \varepsilon_{it}$$

$$\Delta HP_{it} + HP_{i,t-1} = \lambda_i HP_{i,t-1} + \beta_{10i} RGDP_{it} + \beta_{11i} RGDP_{i,t-1} + \mu_i + \varepsilon_{it}$$

$$\Delta HP_{it} = -(1 - \lambda_i) HP_{i,t-1} + \beta_{10i} RGDP_{it} + \beta_{11i}(RGDP_{it} - \Delta RGDP_{it}) + \mu_i + \varepsilon_{it}$$

$$\Delta HP_{it} = -(1 - \lambda_i)(HP_{i,t-1} - \tfrac{\mu_i}{1-\lambda_i} - \tfrac{\beta_{10i}+\beta_{11i}}{1-\lambda_i} RGDP_{it}) - \beta_{11i} \Delta RGDP_{it} + \varepsilon_{it}$$

The same way, using several substitutions, we get the error correction re-parameterization of (6):

$$\begin{aligned} \Delta HP_{it} = \phi_i(HP_{i,t-1} - \theta_{0i} - \theta_{1i} RGDP_{it} - \theta_{2i} HICP_{it} - \theta_{3i} UNEMP_{it}) - \\ - \beta_{11i} \Delta RGDP_{it} - \beta_{21i} \Delta HICP_{it} - \beta_{31i} \Delta UNEMP_{it} + \varepsilon_{it}, \end{aligned}$$

where $\phi_i = -(1 - \lambda_i)$, $\theta_{0i} = \frac{\mu_i}{1-\lambda_i}$, $\theta_{1i} = \frac{\beta_{10i}+\beta_{11i}}{1-\lambda_i}$, $\theta_{2i} = \frac{\beta_{20i}+\beta_{21i}}{1-\lambda_i}$, and $\theta_{3i} = \frac{\beta_{30i}+\beta_{31i}}{1-\lambda_i}$.

The most relevant for us are the error-correction speed of adjustment parameter, ϕ_i, and the long run coefficients θ_{1i}, θ_{2i}, and θ_{3i}. θ_{0i} is the intercept. A negative ϕ_i indicates that house prices tend to return to a long-term equilibrium; the more negative it is, the faster the equilibrium is reached after a deviation.

4.2.3 Robustness

We test for cointegration and unit root. The former needs As Kholodilin *et al.* (2007) notes, there is no test for cointegration in the PMG model, but we can assume cointegration if the autocorrection terms are negative and significantly different from zero, implying that the prices tend to long-term equilibrium. In our case, the autocorrection terms are negative in 9 out of 10 countries. 7 of these 9 are significant on 10% level and and 5 of them even on 1% level of significance. Therefore, similarly to Kholodilin *et al.* (2007) or Égert & Mihaljek (2007), based on we consider the variables cointegrated.

Testing for unit root will be done in two steps. Strauss & Yigit (2003) and Jönsson (2005) in Kholodilin *et al.* (2007) found that conventional unit root tests (DF, PP, KPSS) have a low power in presence of cross section dependence. Therefore we will first test for cross section dependence and then choose a proper unit root tests according to the results.

Table B1 with the results of Pesaran's CD test shows that the null hypothesis of cross section independence was rejected in all cases. Pesaran (1997) or Moon & Perron (2004) suggested unit root tests designed for data with cross section dependence. They have a null hypothesis of nonstationarity of all panels. Hadri *et al.* (2009) further updated the model from Hadri (2000) to allow for cross section dependence, with a useful null hypothesis of stationarity of all panels. Nevertheless, neither the Moon and Perron's test, nor Hadri's tests are currently available for Stata. Therefore we only used the applicable tests: Pesaran's CADF test, and Fisher's unit root test with subtracted cross section mean, which should mitigate the impact of cross section dependence.(Levin *et al.* (2002)) The results can be found in Table B2 and B3. They suggest that before differencing, all the time series were nonstationary but differencing rendered some of them stationary. Unfortunately, we cannot decide whether some time series remained nonstationary even after differencing.

Table 1: Results of Pooled Mean Group estimation

Variable	***Coefficient***	***Std.Error***
Long-run coefficients		
RGDP	2.324***	0.3214
HICP	-0.131	0.1957
UNEMP	-0.032***	0.0098
Short-run coefficients		
RGDP	1.3737***	0.4831
HICP	0.0016	0.2344
UNEMP	-0.0011	0.0059
Adjustment coef.	-0.233513	0.0605

4.2.4 Results

The results of the PMG regression can be found in Table B4. They confirm that both real GDP and - to a lesser extent - unemployment play a role in determining house prices in the long run. They both have the expected sign and are significant at 1% level. Coefficient is quite small in the case of unemployment, however, showing that the link is persistent but weak. On the other hand, the effect of inflation on house prices, suggested by Poterba (1984), did not show: the coefficient was low and insignificant. That could be caused by the relatively low inflation in most countries over the period and by low housing tax subsidy, which would be understandable because of already high owner occupation in most countries.

The adjustment coefficient was significant and negative, which means that in the long run house prices tend to equilibrium. The speed of convergence to equilibrium is relatively slow: it takes over 3 quarters of a year till the deviation decreases to one half of its original magnitude. (half-time=$\frac{ln(0.5)}{ln(1-\phi)}$)

4.3 House price leader effect of national capitals

The results of the analyses based on house price "fundamentals" have some explanatory power but as noted, they must be taken with caution for a number of reasons, from the quality of data (short time series, differing methodologies of data collection, differences between housing stocks, etc.) to the characteristics of data (nonstationarity). These inevitable issues decrease the possibility to draw conclusions and forecast house price developments from the development of fundamentals. Fortunately, there is one more indicator, which can be used to forecast developments in the whole country. It is based on empirical observations, rather than on theory, but that does not diminish its usefulness. We are speaking of the role of national capitals as price leaders.

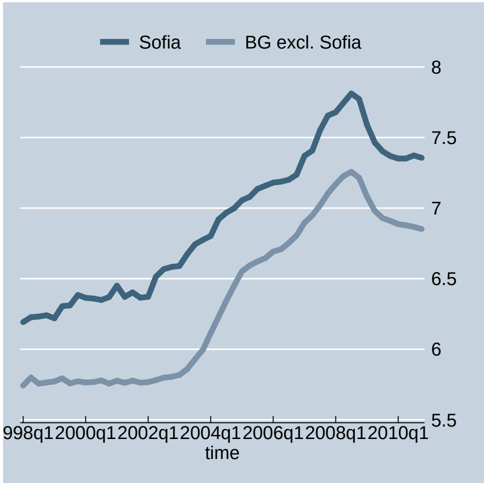
As shown in Hlaváček & Komárek (2009), the lagged correlation coefficient pointed to a certain price leader effect in the Czech Republic, where Prague, lagged by 1 or 2 quarters, would lead the rest of the CR.[20] Our objective here will be to assess whether developments in other countries follow a similar pattern. forecasting house price developments in the short run would be greatly simplified by this, as it would be sufficient to look at the development in the capital to tell with a reasonable likelihood how the prices will change elsewhere. The only pitfall here may be the delay of data publishing; only with up-to-date data would we be able to really *forecast* the development. If the house price data were only published 2 quarters later and the lag was around 2 quarters or less, it could hardly help us. Fortunately, the house price data are mostly published about 1 quarter later, still giving some information about the development in the next quarter.

An advantage of this kind of analysis in comparison with one based on fundamentals is that it should work better when applied on the highly volatile, deviating time series of the crisis. As mentioned at the very beginning of Section 4, the links between fundamentals and house prices often detach during the crisis. In the case of price leader effect, however, it is still likely that the link between prices in the capital and the rest of country would be retained.

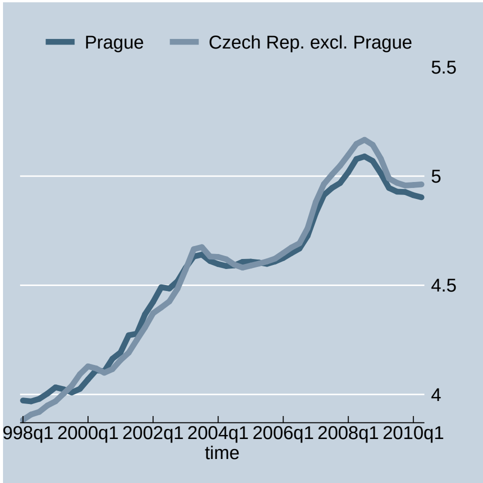
The analysis was performed in all countries where data were available separately for the capital and the rest of the country, or the whole country; the former is naturally preferred but the overall impact of the capital on the price level in the whole country is usually not so significant that it would distort the results. The countries where data was available for the capital and the rest of the country are the Czech Republic, Estonia, and Slovenia. The countries where we had data for the capital and the whole country are Bulgaria and Slovakia. The developments can be compared in Figure 11.

[20] The correlations were significant also between the lagged rest of the CR and Prague, but in most cases (4 house price data sets were used) they were lower. The price leader effect therefore does not stem from the high correlation (the time series are highly correlated almost "inherently") but from the difference between lagged correlations in the opposite "directions"

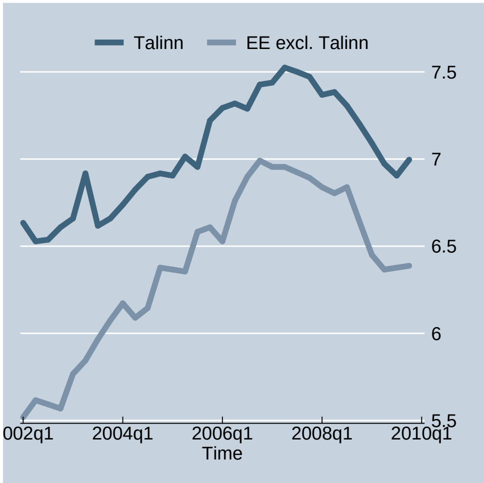
Figure 11: Price leader effect: logged prices in capital and rest of the country



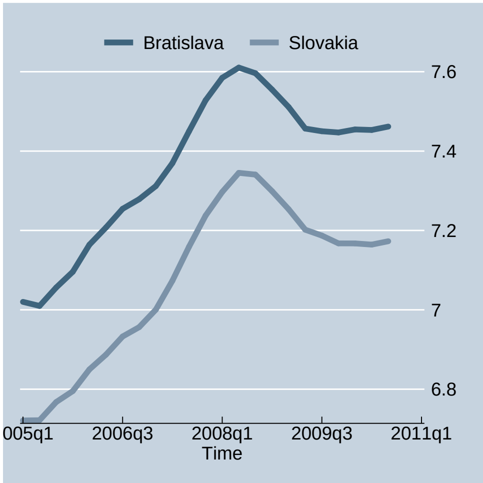
(a) Bulgaria



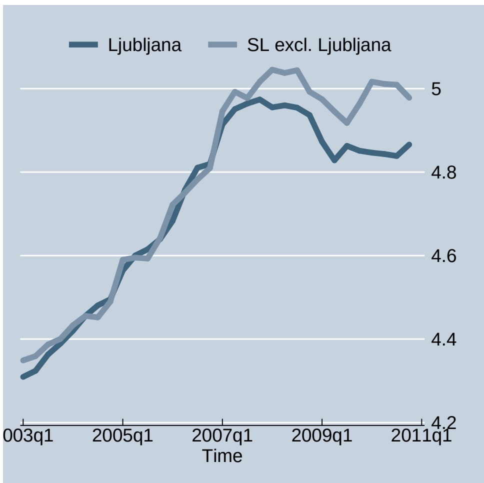
(b) Czech Republic



(c) Estonia



(d) Slovakia



(e) Slovenia

Sources: BIS, national statistical offices, national banks, author's calculations

4.3.1 Model specification

In all the countries, we use the price of houses in the capital and in the rest of the country (or the whole country - but for simplicity we will call both "the rest") and try to estimate the long-term relationship. We expect the capital to be a price leader of the rest, thus being strongly correlated with the "forwarded" (lagged into future) whole-country time series. In most cases we employed the vector autoregressive (VAR) model. We also attempted to use a simple error-correction model from Subsection on page 45, adjusted for this case, but found out it was unsuitable - the time series were not cointegrated or the results were unconvincing.

The VAR model is convenient "*when we are not confident that a variable is actually exogenous*." (Enders (1995)) In this model we treat each variable symmetrically. Its key feature is that the realizations of variable y_t is affected by its past realizations ($y_{t-i}\ i \in \mathbb{N}$) or also by current and past realizations of different variables (e.g. $z_{t-i},\ i \in \mathbb{Z}_0^+$). The length of the longest lag used in the model determines the *order* of the VAR model, marked VAR(n). A simple example of a bivariate VAR(1) model from Enders (1995) is

$$y_t = \beta_{10} - \beta_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt}$$
$$z_t = \beta_{20} - \beta_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \tag{7}$$

Coefficient β_{12} reflects the contemporaneous effect of z_t on y_t. If it is nonzero, any changes in ε_{zt} (e.g. shocks) have an indirect impact on y_t.

The form (7) is called structural VAR. There also exists an expression of VAR in standard form, which expresses the same relationships but is more useful because it shows the current variables as functions of only its past values and the past values of the other variables so it cannot happen as in (7) that in one equation, z_t has an effect on y_t while in the other y_t has an effect on z_t. Nonetheless, the standard form is a little bit more complicated so for illustration the structural VAR should suffice.(for more details, see Enders (1995).

The assumptions of the VAR model are that i) all variables are stationary, ii) all ε_{vt}, (v marks a variable) are nonautocorrelated white-noise disturbances. During the estimation we always check whether i) is fulfilled with unit root tests and test for ii) using the Q-test for autocorrelation of disturbances and a test of normality, e.g. Jarque-Bera or Shapiro-Wilk test. Assumption ii) is analyzed in Subsubsection Robustness.

The exogeneity in a VAR model relaxes the somewhat unrealistic restriction by the linear regression model, which divides the variables between endogenous and exogenous with no interaction between them. In the real world, the macroeconomic quantities are often interrelated so it hardly ever happens that one can say that A influences B but B has no effect - even indirect - on A. In the ability to capture these relations lies indeed an unquestionable advantage of the VAR model. On the other hand, with all variables exogenous and past observations influencing the present one, it may be difficult to distinguish and evaluate the effect of other variables on the one in question. How strongly are A and B connected? Which one has a larger impact on the other than vice versa? And most importantly - does A cause B or B causes A?

The last question is often a difficult one and the VAR model itself does not give us a satisfactory answer. Nonetheless, there exists a method to evaluate which variable really influences which. Although it does in no way give a definite answer to the question of "causality" in the common sense, it allows us to find the answer in the "econometric sense" - can we explain A only from its past observations, or does B also have a significant explanatory power? Based on this question so-called *Granger causality* is defined as a crutch to find the approximate implications between the variables. Even though the conclusions can be economically meaningless (e.g. higher prices of computers could "Granger cause" total inflation), they give us a certain guidance in terms of how important is one variable in explaining an another. In our case, we would like to find out whether house price developments in the capital are a significant "predictor" of the prices in the rest of the country; establishing a Granger causality would support this hypothesis. Therefore we will perform a test for Granger causality after every estimation of a VAR model.

We define Granger causality according to Granger (1969):

Let A_t and B_t be stationary stochastic processes. We denote $P(A|B)$ the optimum, unbiased, least squares predictor of A_t by values of B_t as $P_t(A|B)$, and denote the variance of predictive errors (residuals) of this predictor $\sigma^2(A|B)$. We further denote the set of all past values of a process $\overline{A}$. Further let U_t be all the information in the universe accumulated since time $t-1$, let X_t and Y_t be specified time series, and let $\overline{U_t - Y_t}$ denote the complement of Y_t in U_t.

If $\sigma^2(X|U) < \sigma^2(X|\overline{U-Y})$, we say that Y is causing X, denoted $Y_t \Rightarrow X_t$. We say that Y_t *is Granger causing* X_t if we are better able to predict X_t using all available information than if we only used all the information apart from Y_t.

In praxis, we naturally do not have all the information in the universe, but we can denote all the information we have by U_t and then use Granger causality tests with confidence intervals to find out whether the causality $Y_t \Rightarrow X_t$ is significant. During testing, we have to bear on mind the limited sense of this causality as explained above.

The error-correction model with variables lagged by 1 can be expressed in this form:

$\Delta p_t = \gamma(\Delta c_t) + \lambda(p_{t-1} - \theta c_{t-1}) + \varepsilon_{pt}$

$\Delta c_t = \gamma(\Delta p_t) + \lambda(p_{t-1} - \theta c_{t-1}) + \varepsilon_{ct}$,

where c_t is the house price in the national capital, p_t is the price in the rest of the country, ε_{pt} and ε_{ct} are the disturbance terms in the individual equations.

We estimate all the countries where data are available separately for the capital and the rest of the country (or the country as a whole - naturally there is the problem that such data includes the capital, but we can partially account for this by a more careful interpretation of the results). These countries are Bulgaria, the Czech Republic, Estonia, Slovakia, and Slovenia.

4.3.2 Estimation

In all countries, we used natural logarithm of the price or the price index. The issue of real/nominal prices was irrelevant here, as we were only trying to explain one price development by another: the only important thing was that the figures be comparable. As the source of data and units were the same, this was fulfilled.

The estimation was performed in compliance with the procedure suggested by Cipra (2008). Some concrete steps were inspired by Anton Parlow.[21] We used time series with prices in natural logarithm. We performed tests to find the more suitable from the two models, VAR and VECM. For stationary data, we use VAR model, for nonstationary and cointegrated data we used VECM model. If the data were nonstationary but not cointegrated, we performed none of the tests.

Now let us describe the procedure more in detail (see also Figure 12):

First, we determine the number of lags using Schwert's rule of thumb:

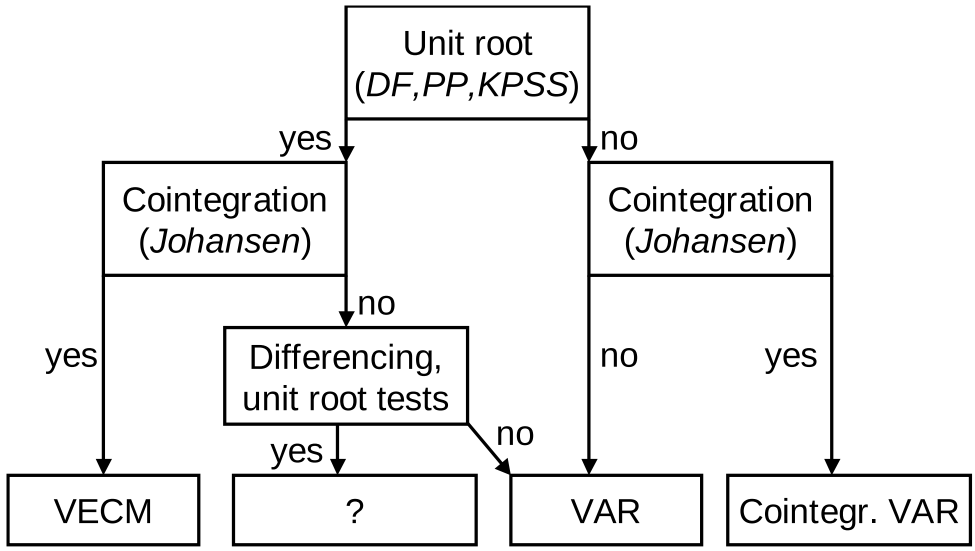
$l_{max} = integer\ part\ of\ \sqrt[4]{12{\times}(T/100)}$, where T is the number of observations.

We perform unit root tests, DF, PP and KPSS, on both of the logged time series, using the calculated number of lags (KPSS calculates the number automatically). Then

[21] https://pantherfile.uwm.edu/aparlow/www/teaching.html

we estimate a VAR model with differenced time series and run the `varsoc` test, based on several information criteria, with l_{max} maximum number of lags to find the optimum lag length, which is a necessary to know in order to test both VAR and VECM.

Figure 12: Process of choosing a suitable model



Sources: Author, partially based on Cipra (2008)

If there was **no unit root**, we test for cointegration and then use the "normal" or cointegrated VAR model with the lag length chosen by `varsoc`. Then we test for Granger causality with `vargranger`. Finally, we test for stability of the model with the `varstable` command.

If there is **unit root**, we would prefer the VECM model if the variables are also cointegrated. We test for cointegration using Johansen `vecrank` test, using again the optimum number of lags, chosen by `varsoc` during VAR estimation. If there is **no cointegration**, we would perform the unit root tests on *differences* of the logged variables and then either choose a VAR test and continue as in the previous paragraphs, or we perform no test as neither VAR nor VECM can be used for non-cointegrated non-stationary time series. A second differencing would certainly render the series stationary but then, what would be the interpretation of the results? If the change of change in A is 0.2 times the change of change in B, what does it tell us? Therefore performing no test seems like a better option.

If there is **cointegration**, we run the VECM model on logged variables with the `vec` command, using the information from Johansen test (this is done automatically in Stata) and the number of lags from `varsoc.`

This process of estimation places no accent on economic thinking. For instance, in the first step we only choose the optimum lag based on an information criteria test, while common sense suggests that the optimal lag will be 1 or 2 quarters, which corresponds to the findings of Hlaváček & Komárek (2009). The effects of change in fundamentals or exogenous shocks should be first felt in the capital, where house turnover, the volume of investment and activity of agents are generally the highest. Choosing a VAR model with a much longer lag ignores this thinking but on the other hand, the shortest lags have a "chance" to show their significance in the test even if much longer lags are also present.

4.3.3 Results

You can find a summary of the results of the tests in Table C2. Here we provide additional details. For simplicity we denote C the capital and R the rest of the country. Further we denote the (lagged) coefficients corresponding to C explaining R as C=>R etc. [22]

1) Bulgaria

Because of no cointegration, we differenced the variables and used the VAR model, using the number of lags given by the information tests, 1. The results show that the model explains prices in the whole country much better ($R^2 = 0,61$) than prices in the capital ($R^2 = 0,27$). This is partially due to the large fluctuations in prices in the capital. Now we would expect the coefficients of C=>R to be positive and significant, which would indicate the price leader effect.

The significant coefficients of R explaining R or C it seems that if there is a price leader effect, it is the rest that leads the capital. The Granger test confirms this conclusion. The reliability of the results is nevertheless diminished by questionable stationarity of the differenced time series - with differences lagged by 1, both DF and PP reject the null hypothesis of a unit root for C but not for R. (see Table C1) KPSS test points to no unit root - stationarity is not rejected for Sofia on 5% level and for R on 1% level.

Graphical analysis shows, on the other hand, that in Sofia the house price boom started about a year and a half earlier than in the rest of the country; after that they moved nearly synchronically. The development in the capital could therefore be viewed as a predictor of a house price boom in the rest.

[22] Lagged prices in general are denoted "Lq.", for instance "Lq.C". Further, we denote the coefficients of prices in the capital, lagged by q quarters, influencing the prices in the rest, Lq.C=>R. For instance, if the results suggest that the capital is a price leader of the rest by 2 quarters, we would write L2.C=>R.

2) Czech Republic

The time series were nonstationary but Johansen test showed no cointegration so we could not use the VECM model. Therefore we obtained stationary series by differencing and used the VAR model. Contrary to our expectations, the significance of the coefficients suggests that the rest of the CR has a larger impact on the price development in Prague than vice versa. Granger test shows the same result.

Yet the signs of coefficients are doubtful: R=>R and R=>C are all significant but while the coefficients of the first lags are positive, the coefficients of the second lags are negative. That could point to a high volatility: the price trends usually last 2 following quarters and then change. The graph in Figure 11, however, does not support this conclusion.

3) Estonia

No cointegration was diagnosed here either but in contrast with the CR, the results of all 3 unit root tests with the exception of DF test for the whole country without capital pointed to no unit root, therefore we deemed the time series stationary and used the VAR(3) model. The results are not too convincing, especially in comparison with what one might observe in the graph: R^2 is only about 0.3 at both variables. What does Granger test say? It suggests that the price development in the country influences that in Talinn but not reversely. The role of Talinn as the price leader, however it may seem plausible from the graph, is not confirmed by the results of the VAR model. From the graphical analysis, though, it seems that another effect is recognizable - a certain convergence of prices during the price boom.

4) Slovakia

Johansen test shows no cointegration and DF, PP, and KPSS test all do not reject unit root even after differencing. Neither the VECM, nor the VAR model are therefore applicable. These results are probably partially caused by short time series of only 23 observations in each of the two series (2005Q1-2010Q3). Due to this shortness, the results of the tests would probably not be very robust even if the requirements of the VAR or VECM model were fulfilled. The only conclusions we can draw are those from the graph. They show that the prices move almost perfectly in sync; rejection of cointegration seems therefore a bit surprising.

5) Slovenia

The results of PP test pointed to no unit root but the p-values of the DF test of over 0.3, and the results of KPSS test for the capital implied we could not reject the null hypothesis of a unit root. Therefore we continued with the test of cointegration. Johansen test showed integration of order 1, only with a lag of 8 quarters, not less. This length was also supported by the information criteria VAR test. Unfortunately, in such a model 16 degrees of freedom would be "used" by the model, which would leave only $23 - 18 = 5$ degrees of freedom for the estimation for the model. The results would in such case be nearly useless.[23] A tentative estimation and a subsequent stability test confirmed this, with 6 out of 16 eigenvalues of the cointegrating vector out of the unit circle.

Therefore we estimated a VAR model. The only significant coefficients were R=>R and C=>R, both lagged by 2 quarters. The sign of R=>R was -0.55, which would point to a high volatility, but it is also certainly caused by the boom and bust development. On the contrary, the significant coefficient 0.568 of C=>R suggests that Ljubljana is the price leader for the rest with a lag of 2 quarters. Granger test supports this hypthesis. Unfortunately, R^2 is low, only 0.3, therefore the conclusion is not quite foolproof. The graph, however, shows that the effect C=>R indeed exists, but it started working only at the peak of the crisis. That confirms the reflection from the beginning of Section 4: in crisis, the behavior of quantities changes; in this case, the prices detached and Ljubljana started "leading" the rest.

4.3.4 Robustness

Although before using the VAR model we differenced the time series and tested for stationarity in order to be able to employ the model, we also have to apply several post-diagnostical tests. These concern the stability of the model (which holds when all the roots of the estimated autoregression polynomial are inside the unit circle), and autocorrelation in residuals and their normality. The negative impacts of autocorrelated residuals are that the ordinary least squares estimate is not the best linear unbiased estimate, the variance of some slope coefficients can be undervalued (which may lead to model mis-specificaton) and the estimate of variance of residuals also (which inflates the coefficient

[23] The number of degrees lost during estimation is in VAR model equal to $L \times p + 1$ and in VEC model $L \times p + 2$, where L is the number of lags and p is the number of regressors. In both models, one more d.f. is lost in estimation of the intercept, in VEC model yet one more d.f. is lost in estimation of the error-correction term.

of determination, increasing the chance of picking a wrong model). (Cipra (2008)) Normality is not a formal assumption of the model but it is highly desirable because it is the necessary condition of the model to be the best among all unbiased estimates, including nonlinear.

In price leader analysis there is no possibility to change the model's specification, therefore the only reason for these tests in assessing the quality of the model and thus the reliability of the results. The results are acceptable: Figure C3 shows that all the roots indeed lie within the unit circle in all models; according to Q-test in Table C4, the null hypothesis of white noise was rejected in none of the countries and the subsequent test of normality showed that in the Czech Republic and Slovenia, the residuals are also normal.

4.4 VAR/VECM analysis of HP in several individual countries

The insufficiently long or even missing time series on volume of mortgages and other possible determinants in some countries did not allow for including these factors in the PMG model in Subsection 4.2. There was the option of simply omitting these countries with lacking data and testing only the rest, but the structure of data together with non-stationarity (even after differencing) would only allow us to test together countries with widely varying characteristics: looking only at the overlaying house price data, (see Table A2 in the Appendix) it would have been possible to "meaningfully" test only these groups of countries: i) BG, EE, CR, LT, HU ii) BG, EE, HU, PL, CR, LT, SL. It is noteworthy that there is no possibility to test a group of countries from one "area", i.e. either <CZ, HU, PL, SK, and maybe SL> or <EE, LT, LV> or <BG, RO, and maybe HU>. In the author's view, it would make very little sense to test together countries that have economically in common virtually nothing, such as the CR and EE. If we were to test a smaller panel, it would have had to be a panel of similar countries. In such a case (small n), the PMG model would have been improper to use and with regard to non-stationarity, it would be necessary to use the VECM model. Then it would be possible to test together small groups of countries from one "area", but a) these panels would be small ($n < 5$) so the effect of "compound information" would be low, and b) the differing time series lengths would make cutting a significant portion of data, as most panel data models (fixed- and random- effects, for instance) and most tests for unit root do not allow for unbalanced panels. Thus we would have probably lost more information by omitting the observations than we would have gained by testing a panel of data rather than separate countries.

Because of these issues, stemming from highly unbalanced panels - attaining longer time series proved to be an impossible task in many cases - we decided to run separate regressions on several countries with sufficiently long house price series. This also enables us to use a larger variety of regressors because previously, we only had several variables with time series available in full length for all the countries, for which reason we had to omit some of the regressors.

4.4.1 Model specification

We employ the vector error correction model (VECM) because unlike other models it allows for nonstationary data and is designed for cointegrated variables, which is exactly our case, as we show in the Robustness part. The model can be derived from the vector autoregression (VAR) model, which was described in Subsubsection 4.3.1. Derivation of the model is not in scope of this work; we will only describe the logic of the model, its advantages and limitations, and its concrete use here.

As noted earlier, cointegration enables us to distinguish between short-term and long-term relationships. The time paths of cointegrated variables are influenced by their deviation from long-run equilibrium, which means that we must take this influence into account. A major advantage of VECM model over VAR model is that it enables us to capture and model this influence, keeping the effect of short-run dynamics of variables on each other.

Here we introduce the general form of the VECM model:

Suppose that two $I(1)$ variables y_t and z_t are cointegrated and that the cointegrating vector is $(1, -\theta)$. Then all three variables $\Delta y_t = y_t - y_{t-1}$, $\Delta z_t - z_{t-1}$, and $(y_t - \theta z_t)$ are $I(0)$. The error correction model

$\Delta y_t = x_t^T \beta + \gamma(\Delta z_t) + \lambda(y_{t-1} - \theta z_{t-1}) + \varepsilon_t$

describes the variation of y_t around its long-run trend (given by a set of $I(0)$ exogenous factors x_t), the variation of z_t around its long-run trend, and the error correction $(y_t - \theta z_t)$. (Greene (2003)) γ is the short-run coefficient, reflecting the impact of z_t on y_t.

We can divide the model into two parts to get a better understanding of it:

$\Delta y_t = x_t^T \beta + \gamma(\Delta z_t) + \varepsilon_t$ is the equilibrium relationship, and $\lambda(y_{t-1} - \theta z_{t-1})$ is the equilibrium error, which accounts for the deviation of the pair of variables from that equilibrium. The model states that the change in y_t from the previous period consists of the change associated with movement with x_t along the long-run equilibrium path plus a part γ of the shift given by change in Δz_t plus a part λ (speed of adjustment coefficient) of the deviation $(y_{t-1} - \theta z_{t-1})$ from the equilibrium. (Greene (2003)) The coefficient λ points to the speed of reaction of variable y to deviations from equilibrium. The greater λ is, the greater the reaction of y_t after a deviation. (Enders (1995))

4.4.2 Estimation

We analyzed four countries: Bulgaria, the Czech Republic, Estonia, and Lithuania. They were chosen for several reasons: i) they have sufficiently long time series ii) BG, EE, and LT have had similar house price developments; an analysis could help us find out whether they can be explained by the same fundamentals iii) for two of these countries we have house price data in euros per m^2, not just an index. This gives us additional information regarding convergence of the prices together.

Now we turn to the estimation itself. Compared to the previous estimation with the PMG model, we have a larger variety of explanatory variables available. We explain all the variables besides unemployment in natural logarithms. The list of variables is here: real house prices - real GDP, unemployment, HICP, real direct investment, labor force, real volume of mortgages per person, real construction price index. The data are more precisely described in Table A1.

First we tested whether price developments could be explained only by their past values. Is it possible? The boom and bust of the crisis suggests that it should not be possible. The estimation was performed as follows:

At the beginning of the autoregression we found the optimal lag length by `varsoc`, then performed tests for unit root. If there was none, we ran a VAR model with the chosen lag length. In the other case we used shorter lags and tested for unit root again, until the test showed no unit root or until the lag would be zero. In that case we could not use the VAR model. The results can be found in the Appendix in Table D1. They show that only in Bulgaria it would be well possible to explain prices based on their past values; all the other models had R^2 of around 0.1.

This result can only be thought of as a measure of how much the model could be improved by adding other variables. For example, if R^2 of a model $P_t = \alpha + \beta_1 P_{t-1} +$

$\beta_2 P_{t-2} + \varepsilon_t$ was 0.8, it would mean that there would be little point adding new variables. This could only be avoided by excluding these past variables of house prices from the model. In VAR or VECM this is not possible but with regards to the results, except for Bulgaria there would be no need to omit the past values of house prices.

As mentioned, this impossibility to find a good autoregressive model for Estonia and Lithuania could be attributed to the boom and bust dynamics of the crisis; therefore we tried to run the same test only with the data before the price peak. In Lithuania, the model did not improve at all, but in Estonia it improved significantly: the coefficient of determination nearly reached 0.5. Nonetheless, the coefficients at P_{t-1} and P_{t-2} were -0.8 and -0.54 respectively, only showing a high volatitility.

After the "pure autoregression" model, we tried to find the best model specification for each country. The final model should

a) be "simple" enough not to use too many degrees of freedom on estimates of many variables and their past values, thus spoiling the results

b) have significant explanatory variables (of house prices) and a high coefficient of determination

c) be reasonably stable.

The estimation was very similar to the one in 4.3.2 (see also Figure 12) and consisted of running uncountable tests for unit root, cointegration, and information criteria (for assumptions of VAR or VECM to be fulfilled), and trying different model specifications that would both be consistent with these tests and yield satisfactory results. Our search lead in each of the countries to a different result. This seems surprising especially in case of Lithuania and Estonia, two countries with very similar characteristics.

We always first attempted to use the VEC model, which better captures long term relations between variables (Cipra (2008)), but sometimes it proved impossible to find a set of variables that would be cointegrated and at the same time yield satisfactory results. In that case, we differenced the variables and if the series were stationary, we applied the VAR model.

The final specifications were: $HP \sim HICP + LABF + MORTG$ for Bulgaria, $HP \sim GDP + UN$ for the Czech Republic and Lithuania, and $HP \sim GDP + DIRINV$ for Estonia.

4.4.3 Results

The results are summed up in Table 2 and in more detail in Tables D2 and D3 in the Appendix. It must be taken into account that all of the variables that were not included in the final chosen model specification showed no significance; the value of this analysis therefore lies not only in suggesting which determinants have a significant impact on house price developments in the respective country, but also in showing which have not.[24]

Table 2: Results of VAR/VECM analysis

		VECM	VECM	VAR	VAR
		BG	*CZ*	*EE*	*LT*
lags()		(2)	(1)	(2)	(2)
error-corr. term		-0.24***	0.006	-	-
HP	D1	0.606***	0.285*	-0.589***	-0.068
	D2	0.384*		-0.301**	-0.059
GDP	D1		-1.772***	3.242***	0.97*
	D2			2.057	-0.145
UN	D1		-0.019		-0.006
	D2				-0.032***
HICP	D1	-0.248			
	D2	-0.291*			
LABF	D1	-0.069			
	D2	0.339**			
MORTG	D1	0.052			
	D2	-0.004			
DIRINV	D1			-0.065***	
	D2			-0.03	
R2		0.85	0.59	0.45	0.36

1) Bulgaria

Because of nonstationarity of differenced house prices we used the VEC model, even though the "preferable" sets of time series were sometimes not cointegrated, which limited the choice of variables. Of all the determinants tested, only HICP and labor force turned out to be significant. Unfortunately, these time series together with house prices were not cointegrated. We circumvented this obstacle by adding the volume of mortgages to the equation: now, the series were cointegrated, which allowed for using VECM at a cost of several degrees of freedom used to calculate the coefficients for mortgages.

Both the significant coefficients and the "Granger test" suggest that HICP and labor force have an impact on house prices with a lag of two quarters. Surprisingly, the signs

[24]Again, the list of tested variables is: real GDP, unemployment, HICP, real direct investment, labor force, real volume of mortgages per person, real construction price index.

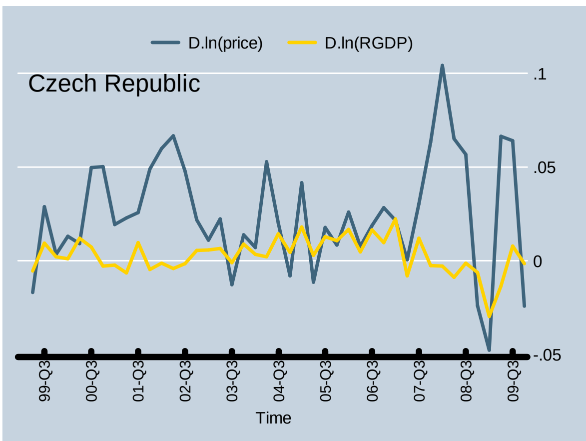
of the coefficients at HICP are negative. As Iacoviello (2000) notes, higher inflation rates reduce homeowners' costs because the capital gains from house appreciation are untaxed. Therefore a transitory inflation shock should increase demand for housing and thus drive house prices up. On the other hand, *"inflation increases also drive endogenous movements in output*[25] *and and interest rate that might counterbalance the effect."* Iacoviello (2000) His empirical results show that house prices initially dropped after an inflation shock in 3 out of 6 analyzed countries.[26] These drops demonstrated themselves within one or two lags after the shock, which corresponds with our results of negative signs of HICP coefficient with a 1- and 2-quarters lag.

The positive effect of changes in labor force with a lag of two quarters is quite predictable, as it reflects the increased demand after housing. And lastly, the error correction term is significant and negative, which means that in the long run house prices return to equilibrium. The half-time of the return is a little over 3 quarters.[27]

2) Czech Republic

Here, our "optimum model" uses more conventional determinants: real GDP and unemployment. The former is highly significant but it has a negative coefficient, from which we assume that the model must be flawed because probably all studies of house prices show that the correlation between real GDP and house prices is positive. Let us at least analyze the data graphically in order to find a clue.

Figure 13: Comparison of differenced logs of real house prices and real GDP in the Czech Republic



[25]If output drives house prices, a drop in house prices would lead to a decrease in house prices.
[26]These countries were Germany, Italy, and the UK.
[27]See the end of PMG analysis: $half - time = \frac{ln(0.5)}{ln(1-\phi)} = \frac{ln(0.5)}{ln(1-(-0.24))} \doteq -3.2$

From Figure 13 we can partially derive the reason why the coefficient of GDP is so significantly negative: even if not always, the time series largely diverge. We can further make use of this graph by taking real GDP as a proxy for equilibrium house price development (which is not unrealistic, given the unquestionable link between GDP and house prices), we see that house prices deviated largely between 2001 and 2002, and again between the end of 2007 and half of 2008. This corresponds to the results of Hlaváček & Komárek (2009), who found a property price bubble in 2002/2003 and a potential bubble in 2007/2008.[28] A reason for the difficulty of finding a suitable model can therefore lie simply in factors we are unable to measure - expectations, exogenous shocks, etc.

3) Estonia

Because we were unable to find a set of variables that would both be cointegrated and yield acceptable results in VECM, we used a VAR model. Real GDP and direct investment together with the past values turned out to be the only significant variables. The negative signs of lagged HP point to a large volatility in the time series, which is by all means in compliance with our expectations based on Figure 5. Real GDP, unlike in the Czech Republic, is confirmed as a significant driver of house prices; according to the coefficient, if GDP grew by 1 percent per year, house prices would ceteris paribus grow by 3.24%! Direct investment was expected to push house prices up, too, as it increases demand for labor, which increases wages, which increases demand for housing, which would increase house prices. On the other hand, the direct investment includes construction of residential buildings; as we can see in Figure 9, although Estonia was no "construction tiger", the increase in number of houses together with decreasing population made the market more saturated, which could lead to lower prices. In every case, the magnitude of the coefficient is low so direct investment is by far not as important as real GDP.

4) Lithuania

As in Estonia, also here seems GDP to play a role, even though the significance and magnitude of the coefficient and the low R^2suggest that the role is less important. Yet this result may be caused by the inclusion of unemployment in the model - even though the variables were not cointegrated, it is a well-known fact that GDP and unemployment are strongly negatively correlated. Unemployment could have therefore "replaced" GDP as an explanatory variable. But even if this was true, the coefficient is very low - not

[28]The price increase was comparable according to their results but in 2007/2008 it was explainable by the underlying fundamentals. In Figure 9 we see that real GDP was hardly one of these fundamentals.

enough for unemployment, which changed only in order of percentage points, to explain house price surges of tens of percent a year (see Figure 2). R^2 is also quite low; it seems that other factors besides the ones we tested were present. This comes as no surprise as small housing markets, such as those in the Baltic states, are highly volatile and prone to swift, hardly explainable and unpredictable property price movements.

For illustration, we also created Figure D4, which shows how volatile real house prices have been in comparison to real GDP, and therefore how difficult it is to predict price developments based on it. The graphs show that small changes in GDP are sometimes met with large surges of house prices and that these surges are quite unpredictable.

5 Conclusion

The goal of this work was to find the most important factors that have influenced house price in ten countries of Central and Eastern Europe over the last decade and discover how and how fast the prices react to deviations from long-term equilibrium. We used several methods of analysis: graphical comparison of the characteristics of the housing markets, Pooled Mean Group analysis of panel data, and VAR or VECM analysis of time series.

House price developments are influenced by the conditions on housing markets. Therefore we analyzed and compared housing market characteristics in the individual countries, searching for a link with house price developments. Some connections were found: owner occupation and net migration seem to play a role. On the other hand, several other characteristics seem not to have had any clear impact. For instance, the nearly three times higher volume of mortgages per capita in Estonia than in Latvia made little difference in price developments in these countries.

During the analysis of house price determinants we first used the Pooled Mean Group (PMG) model on a panel of all countries, explaining real house prices in euros by three variables: real GDP, HICP, and unemployment. Both real GDP and unemployment proved significant on 1% level, HICP was not significant. The half-time of return of house prices back to long run equilibrium was in average quite slow: over 3 quarters.

A house price analysis based on fundamentals requires that relationships between house prices and fundamentals be stable. If these change, for instance in crises, models cease working. Performing a "price leader analysis", which assumes that prices in the country follow with a delay its capital, the "price leader". The availability of data enabled us to perform this analysis in four countries, using the vector autoregressive (VAR) model. To establish causality, we analyzed significances and used the Granger test. The price leader effect was found in Bulgaria, where Sofia leads the rest by 4 quarters, and in Slovenia, where Ljubljana leads the rest by 3 quarters.

In the next section, we employed a vector error-correction model (VECM) or a VAR model on a wider set of explanatory variables in four countries with sufficiently long time series and attempted to find the best model specifications. The results vary across countries. For instance, in Bulgaria the past values alone were able to explain over 60%

of house price variance, contrary to under 20% in other countries. In most countries, GDP played a major role but usually there was one more important, country-specific determinant.

The overall results suggest that the determining factors vary across the countries, as do the price development patterns. The policy makers should therefore bear these country-specific factors on mind. The heterogeneous results may however also be partially attributable to the unavoidable issues of variable house price data quality and short time series, which limit comparability of the results. Yet these issues are likely to ameliorate in the future - as new countries (e.g. Romania) started publishing house price indices, more data will be available, and the methodology may unite based on the guide for developing consistent house price indices by Eurostat. Thus we are confident that house price determinants in CEE, however presently folded in mist, will once become suitable for robust analysis.

References

AHEARNE, A. G., J. AMMER, B. M. DOYLE, L. S. KOLE, & R. F. MARTIN (2005): "Monetary policy and house prices: A cross-country study." *Working Papers Central Bank of Chile 344*, Central Bank of Chile.

ALLEN, F. & E. CARLETTI (2010): "Financial regulation going forward." *IMES Discussion Paper Series 10-E-18*, Institute for Monetary and Economic Studies, Bank of Japan.

ANNETT, A. (2005): "House prices and monetary policy in the euro area." *IMF Country Report 05/266*, International Monetary Fund.

BALTAGI, B. H. (2005): *Econometric Analysis of Panel Data.* John Wiley and Sons.

BREITENFELLNER, B. & N. WAGNER (2010): "Government intervention in response to the subprime financial crisis: The good into the pot, the bad into the crop." *International Review of Financial Analysis* **19**: pp. 289–297.

CAHLÍK, T., M. HLAVÁČEK, & J. SEIDLER (2010): *Makroekonomie.* Karolinum.

CAMPBELL, J. Y. & J. F. COCCO (2007): "How do house prices affect consumption? evidence from micro data." *Journal of Monetary Economics* **54(3)**: pp. 591–621.

CHEN, J. (2006): "The dynamics of housing allowance claims in sweden: A discrete-time hazard analysis." *Working Paper Series 2006:1*, Uppsala University, Department of Economics.

CIPRA, T. (2008): *Finanční ekonometrie.* Ekopress.

COLANDER, D., H. FOLLMER, A. HAAS, M. GOLDBERG, K. JUSELIUS, A. KIRMAN, T. LUX, & B. SLOTH (2009): "The financial crisis and the systemic failure of academic economics." *Kiel Working Papers 1489*, Kiel Institute for the World Economy.

EDELSTEIN, R. H. & S. K. LUM (2004): "House prices, wealth effects, and the singapore macroeconomy." *Journal of Housing Economics* **13(4)**: pp. 342–367.

ENDERS, W. (1995): *Applied Econometric Time Series.* Wiley.

EUROSTAT (2011): "Draft version 3.0 of handbook on residential property price indices." *Technical report*, European Commission.

ÉGERT, B. & D. MIHALJEK (2007): "Determinants of house prices in central and eastern europe." *BIS Working Papers 236*, Bank for International Settlements.

Girouard, N., M. Kennedy, P. van den Noord, & C. André (2006): "Recent house price developments: The role of fundamentals." *OECD Economics Department Working Papers 475*, OECD Publishing.

Granger, C. W. J. (1969): "Investigating causal relations by econometric models and cross-spectral methods." *Econometrica* **37(3)**: pp. 424–38.

Greene, W. H. (2003): *Econometric Analysis.* Prentice Hall.

Hadri, K. (2000): "Testing for stationarity in heterogeneous panel data." *Econometrics Journal* **3(2)**: pp. 148–161.

Hadri, K., S. de Silva, & A. R. Tremayne (2009): "Panel unit root tests in the presence of cross-sectional dependence: finite sample performance and an application." *Econometrics Journal* **12(2)**: pp. 340–366.

Hilbers, P., A. W. Hoffmaister, A. Banerji, & H. Shi (2008): "House price developments in europe: A comparison." *IMF Working Paper 08/211*, International Monetary Fund.

Hlaváček, M. & L. Komárek (2009): "Housing price bubbles and their determinants in the czech republic and its regions." **(12/2009)**.

Iacoviello, M. M. (2000): "House prices and the macroeconomy in europe: results from a structural var analysis." *Working Paper Series 18*, European Central Bank.

Jönsson, K. (2005): "Cross-sectional dependency and size distortion in a small-sample homogeneous panel data unit root test." *Oxford Bulletin of Economics and Statistics* **67(3)**: pp. 369–392.

Kholodilin, K. A., J.-O. Menz, & B. Siliverstovs (2007): "What drives housing prices down?: Evidence from an international panel." *Discussion Papers of DIW Berlin 758*, DIW Berlin, German Institute for Economic Research.

Levin, A., C.-F. Lin, & C.-S. James Chu (2002): "Unit root tests in panel data: asymptotic and finite-sample properties." *Journal of Econometrics* **108(1)**: pp. 1–24.

Miles, D. (1994): *Housing, Financial Markets and the Wider Economy.* Wiley.

Moon, H. R. & B. Perron (2004): "Testing for a unit root in panels with dynamic factors." *Journal of Econometrics* **122**: pp. 81–126.

Muellbauer, J. & A. Murphy (1997): "Booms and busts in the uk housing market." *CEPR Discussion Papers 1615*, C.E.P.R. Discussion Papers.

Pagés, J. M. & L. A. Maza (2003): "Analysis of house prices in spain." **(0307)**.

Palacin, J. & R. Shelburne (2005): "The private housing market in eastern europe and the cis." *ECE Discussion Papers Series 2005/5*, UNECE.

Pesaran, H. (1997): "Pooled estimation of long-run relationships in dynamic heterogeneous panels." **(9721)**.

Poterba, J. M. (1984): "Tax subsidies to owner-occupied housing: An asset-market approach." *The Quarterly Journal of Economics* **99(4)**: pp. 729–52.

Ratcliffe, A. (2010): "Housing wealth or economic climate: Why do house prices matter for well-being?" *The Centre for Market and Public Organisation 10/234*, Department of Economics, University of Bristol, UK.

Stepanyan, V., T. Poghosyan, & A. Bibolov (2010): "House price determinants in selected countries of the former soviet union." *IMF Working Papers 10/104*, International Monetary Fund.

Strauss, J. & T. Yigit (2003): "Shortfalls of panel unit root testing." *Economics Letters* **81**: pp. 309–313.

Sutton, G. D. (2002): "Explaining changes in house prices." *BIS Quarterly Review September*.

Terrones, M. & C. Otrok (2004): "The global house price boom." *World Economic Outlook April*, International Monetary Fund, Chapter II in IMF.

Tsatsaronis, K. & H. Zhu (2004): "What drives housing price dynamics: cross-country evidence." *BIS Quarterly Review* .

Weber, A. A. (2008): "Financial market stability." .

List of appendices

A Appendix

DATA

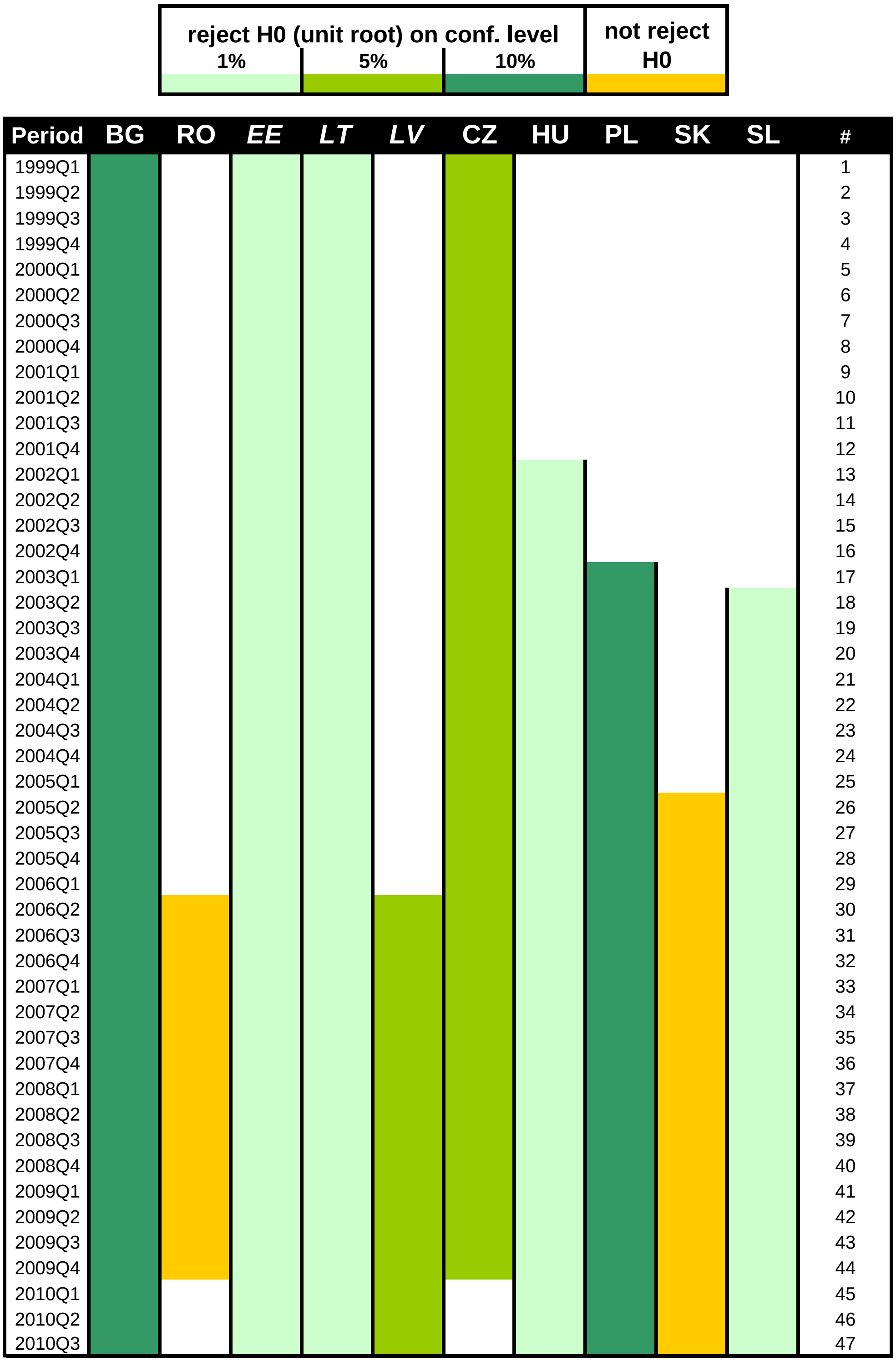
Table A1) Data description

Variable	Code	Notes	Frequency	Type	Unit	First data available for all	Last data	Source	Code in the source
Real GDP	RGDP	SA	Q	chain-linked volumes, reference year	millions EUR	1997Q1 9 countries, 2000Q1 all	2010Q3	Eurostat	namq_gdp_k
Unemployment	UNEMP	From 25 to 74 years	Q	-	percent	2001Q4	2010Q2	Eurostat	lfsq_urgan
Inflation	HICP		M => Q	HICP (2005 = 100)	-	1997M1	2010M12	Eurostat	prc_hicp_midx
Labor force	LABF		Q	sum	persons in 1000s	ca 2001Q2	2010Q2	IMF IFS	9****D..ZF...
Construction cost index of new residential buildings	CONS		Q	index (2005=100)	-	2000Q1	2010Q3	Eurostat	sts_copi_q
Direct investment	DIRINV		Q	volumes	millions USD	2000Q1	2010Q1	IMF IFS	9***8BEDZF...
House purchase loans	MORTG		M => Q	volumes	millions EUR	2004M1	2010M11	EMF	-

Country	First data	Last data	Coverage	Source	Type of data
Bulgaria	1993Q1	2010Q4	Country ; district centers	National Statistical Institute	average market price per m2 of dwellings
Czech Rep.	1998Q1	2009Q4	Country ; regions	National Statistical Office	quarterly price index and yearly published prices of family houses, flats, residential buildings
Estonia	1997Q1	2009Q5	Capital ; large towns	National Statistical Office	average market price per m2 of flats with 3 rooms and a kitchen
Hungary	2001Q4	2010Q3	Capital	Central Statistical Office	average price per m2 of existing dwellings
Latvia	2006Q1	2010Q2	Country	National Statistical Office	average price per m2 of existing flats
Lithuania	1998Q4	2010Q4	Country	Central Bank	average price per m2 of existing dwellings
Poland	2002Q3	2010Q3*	Capital ; large towns	Central Bank	cost of construction of a m2 in a residential building
Romania	2006Q1	2009Q4	Capital	Real estate agency	asking-price per built square meter of apartments built before 1990
Slovakia	2005Q3	2010Q4	Country ; regions	National Bank	average price per flats and houses
Slovenia	2003Q1	2010Q3	Capital; rest of country	National Statistical Office	price index based on prices of existing flats and family houses

* in Poland from 2002Q3 till 2005Q3 only semi-annual data is available

Table A2) Overview of house price data&its stationarity *in first difference*
The colored lines show available house price data in each country. They were tested first in log, then in first difference by both the Dickey-Fuller and Phillips-Perron test. In log, all had unit root, therefore only results for differenced data are depicted. The p-values of both tests were in all cases in the same "confidence group" (e.g.5%).



Note: the first value displayed in each column was lost due to differencing.

POOLED MEAN GROUP ANALYSIS OF PANEL DATA

Table B1) Results of Pesaran's CD test for cross-section independence of time series

Variable	CD test statistic
logHP	19.38***
logRGDP	25.66***
logHICP	31.51***
UNEMP	16.25***

H_0 : The time series are cross-section independent; *** means rejection of null hypothesis on 1% level of significance.

Table B2) Results of Pesaran's pescadf test for unit root

all variables besides UNEMP are in natural logarithms

	CADF(0)	CADF(1)	CADF(2)
HP	0.244	0.426	0.753
RGDP	0.998	0.996	0.977
HICP	0.685	0.589	0.248
UNEMP	**0**	0.999	0.999

the variables besides UNEMP are differenced natural logarithms

	CADF(0)	CADF(1)	CADF(2)
D.HP	**0**	**0**	**0.005**
D.RGDP	**0**	**0.024**	0.796
D.HICP	**0**	**0**	**0**
D.UNEMP	**0**	**0**	**0.017**

Note: The figures in the tables are p-values. The number of lags is in brackets. H_0 is that all time series are nonstationary.

Table B3) Results of Fisher's xtfisher test for unit root

the variables besides UNEMP are differenced natural logarithms

	lags(0)	lags(1)	lags(2)
D.HP	**0**	**0**	**0.005**
D.RGDP	**0**	**0**	**0.0014**
D.HICP	**0**	**0**	**0**
D.UNEMP	**0**	**0**	**0**

Note: the figures in the tables are p-values. H_0 is that all time series are nonstationary.

Table B4) Detailed results of Pooled Mean Group estimation

Number of obs.:		306
Number of groups:		10
Obs. per group:	min	15
	avg	30.6
	max	41

D.lc	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ec						
lrgdp	2.324147	.3213779	7.23	0.000	1.694258	2.954036
lhicp	-.1308725	.1957143	-0.67	0.504	-.5144654	.2527204
un	-.0320413	.0098217	-3.26	0.001	-.0512914	-.0127912
SR						
ec	-.233516	.0605024	-3.86	0.000	-.3520985	-.1149334
lrgdp D1.	1.373672	.4831166	2.84	0.004	.4267811	2.320563
lhicp D1.	.0015894	.2344976	0.01	0.995	-.4580175	.4611964
un D1.	-.0011152	.0059596	-0.19	0.852	-.0127959	.0105655
_cons	-1.246507	.3207859	-3.89	0.000	-1.875235	-.6177777

PRICE LEADER EFFECT ANALYSIS

Table C1) Tests preceding VAR analysis of price leader effect

Tests of stationarity

	BG	*CZ*	*EE*	*SK*	*SL*
dfuller, lags(x)	(10)	(10)	(9)	(8)	(9)
log(country)	**0.3673**	**0.5074**	**0.3354**	0.0384	**0.3891**
log(cap.)	**0.9421**	**0.4235**	**0.8886**	0.0251	**0.391**
pperron, lags(x)	(10)	(10)	(9)	(8)	(9)
log(country)	**0.8148**	**0.5781**	**0.3043**	**0.4053**	**0.4681**
log(cap.)	**0.9487**	**0.6452**	**0.5492**	**0.3479**	**0.2857**

Tests of stationarity after differencing

	BG	*CZ*	*EE*	*SK*	*SL*
dfuller, lags(x)	(1)	(2)	(2)	(3)	(2)
D.log(country)	**0.257**	0.101	**0.479**	**0.312**	**0.323**
D.log(cap.)	0.019	0.031	0.095	**0.656**	**0.536**
pperron, lags(x)	(1)	(2)	(2)	(3)	(2)
D.log(country)	**0.131**	0.044	0.001	**0.442**	0.0035
D.log(cap.)	0.0003	0.006	0	**0.339**	0.0265

Optimum number of lags (varsoc)

	BG	*CZ*	*EE*	*SK*	*SL*
maxlag(x)	(10)	(10)	(8)	(4)	(8)
FPE	1	2	2	3	2
AIC	1	2	2	3	8
HQIC	1	2	2	3	8
SBIC	1	2	2	1	0
final	**1**	**2**	**2**	**none***	**8**

* no test was eventually run because of unfulfilled requirements

Cointegration (vecrank)

	BG	*CZ*	*EE*	*SK*	*SL*
lags(x)	1; 2	1; 2; 4	2	3	3; 4; 6; 8
maximum rank	0	0	0	-	0; 0; 0; 1

H_0 of DF and PP test of stationarity is unit root.

Table C2) Results of VAR analysis of price leader effect

Results of VAR analysis

		BG	CZ	EE	SL
	lags	1	2	2	2
coef. R=>R	D1	**0.877*** (0.132)**	**0.872*** (0.17)**	0.197 (0.171)	0.079 (0.207)
	D2		**-0.39** (0.179)**	-0.25 (0.181)	**-0.55*** (0.214)**
coef. C=>R	D1	-0.115 (0.112)	0.206 (0.174)	**0.341* (0.178)**	0.379 (0.271)
	D2		-0.02 (0.17)	**0.326** (0.165)**	**0.568** (0.25)**
coef. R=>C	D1	**0.583*** (0.213)**	**0.753*** (0.164)**	**0.397** (0.176)**	0.262 (0.173)
	D2		**-0.561*** (0.172)**	0.197 (0.186)	0.009 (0.179)
coef. C=>C	D1	0.036 (0.181)	0.182 (0.168)	-0.228 (0.183)	0.216 (0.226)
	D2		0.24 (0.164)	0.024 (0.17)	0.115 (0.209)
R^2 R		0.61	0.60	0.25	0.30
R^2 C		0.27	0.56	0.23	0.29

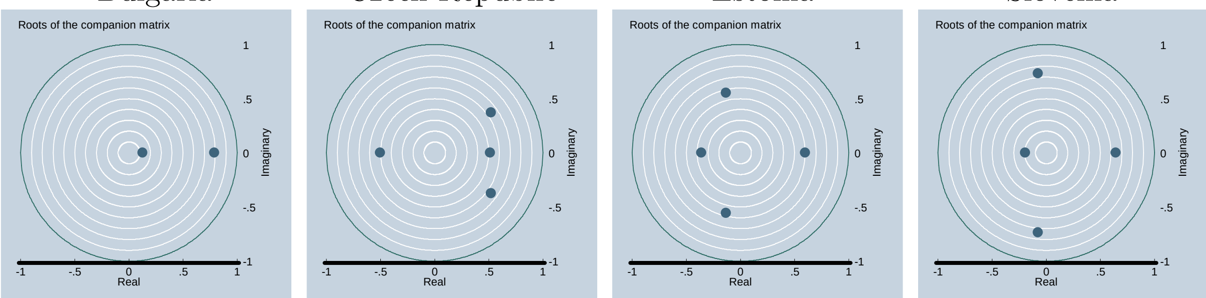
Results of Granger causality test (vargranger)

	BG	CZ	EE	SK	SL
C=>R	0.307	0.494	**0.037**	-	**0.014**
R=>C	**0.006**	**0**	**0.026**	-	0.317

R is the rest of the country, C is the capital, R=>C is the coefficient at R explaining C, etc. R^2 C is the coefficient of determination, corresponding to C, etc. D1 marks first difference of the former time series, lagged by 1, D2 lagged by 2.

Figure C3) Test of stability

Bulgaria | Czech Republic | Estonia | Slovenia



The model is stable (stationary), if all inverse roots of the estimated autoregression polynomial lie within the unit circle.

Table C4) Tests of residuals of VAR models

	BG	CZ	EE	SL
p-value of Q-test	**0.6189**	**0.7002**	**0.2814**	**0.6352**
p-value of Jarque-Bera test	0	**0.6824**	0.037	**0.2706**

In Q-test, H_0 is that the residuals are white noise. In Jarque-Bera test, H_0 is normality of the residuals.

VAR/VECM ANALYSIS OF HOUSE PRICE DETERMINANTS

Table D1) Results of VAR analysis of house prices based only on past values

Optimum number of lags *(varsoc)*

	BG	LT	EE	CZ
maxlag(x)	(8)	(8)	(2)	(4)
FPE	1	4	1	2
AIC	1	4	1	2
HQIC	1	0	1	2
SBIC	1	0	0	0
final	**1**	**1***	**1**	**2**

* a unit root was present with 4 lags

Stationarity

	BG	LT	EE	CZ
dfuller, lags(x)	(1)	(4); (1)	(1)	(2)
D.log(country)	0.255	0.486; 0.002	0	0.003
pperron, lags(x)	(1)	(4); (1)	(1)	(2)
D.log(country)	0.259	0; 0	0.024	0

Results of the VAR model

	BG	LT	EE	CZ
L1.Dlprice	0.82*** (0.087)	0.243* (0.143)	-0.277** (0.148)	0.355** (0.161)
L2.Dlprice				-0.235 (0.16)
R^2	0.666	0.061	0.077	0.118

L1.Dlprice marks differenced natural logarithm of house price, lagged by one, etc.

Table D2) Detailed results of VAR/VECM analysis of house prices

ltotal	ln(real[29] house price or real index)	*ldirinv*	ln(real direct investment)
lhicp	ln(HICP)	*lmortg*	ln(real vol. of mortgages per inhab.)
lgdp	ln(real GDP)	*un*	unemployment in percent

Bulgaria

Vector error-correction model

No. of obs = 41

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_ltotal	10	0.027	0.846	170.33	0

	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
_ce1						
L1.	**-0.24*** **	0.077	-3.1	0.002	-0.3919	-0.0882
ltotal						
LD.	**0.6059*** **	0.168	3.6	0	0.2758	0.9361
L2D.	**0.3842***	0.204	1.88	0.06	-0.0159	0.7842
lhicp						
LD.	-0.2483	0.193	-1.29	0.197	-0.6259	0.1293
L2D.	**-0.2913***	0.177	-1.65	0.1	-0.6380	0.0554
llabf						
LD.	-0.0687	0.156	-0.44	0.659	-0.3735	0.2361
L2D.	**0.3386****	0.156	2.17	0.03	0.0324	0.6448
lmortg						
LD.	0.0519	0.088	0.59	0.554	-0.1198	0.2236
L2D.	-0.0042	0.065	-0.06	0.949	-0.1307	0.1224
_cons	0.0043	0.009	0.48	0.633	-0.0132	0.0217

Czech Republic

Vector error-correction model

No. of obs = 42

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_ltotal	5	0.027	0.587	52.551	0

	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
_ce1						
L1.	**0.0064****	0.005	1.28	0.2	-0.0034	0.0163
ltotal						
LD.	**0.2850***	0.154	1.85	0.064	-0.0169	0.5868
lgdp						
LD.	**-1.772*** **	0.574	-3.09	0.002	-2.8982	-0.6468
un						
LD.	**-0.0187**	0.012	-1.57	0.116	-0.0420	0.0046
_cons	0.0283	0.007	4.1	0	0.0148	0.0419

Estonia

Vector autoregression

No. of obs = 41

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_ltotal	7	0.145	0.446	33.047	0

	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
ltotal						
LD.	**-0.5887*** **	0.152	-3.89	0	-0.8856	-0.2917
L2D.	**-0.3009** **	0.136	-2.21	0.027	-0.5679	-0.0340
lgdp						
LD.	**3.2421*** **	1.244	2.61	0.009	0.8045	5.6796
L2D.	2.0569	1.442	1.43	0.154	-0.7686	4.8825
ldirinv						
LD.	**-0.0649*** **	0.023	-2.82	0.005	-0.1100	-0.0198
L2D.	-0.0304	0.026	-1.17	0.241	-0.0813	0.0205
_cons	0.0113	0.022	0.52	0.604	-0.0313	0.0538

Lithuania

Vector autoregression

No. of obs = 43

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_ltotal	7	0.076	0.381	26.495	0.0002

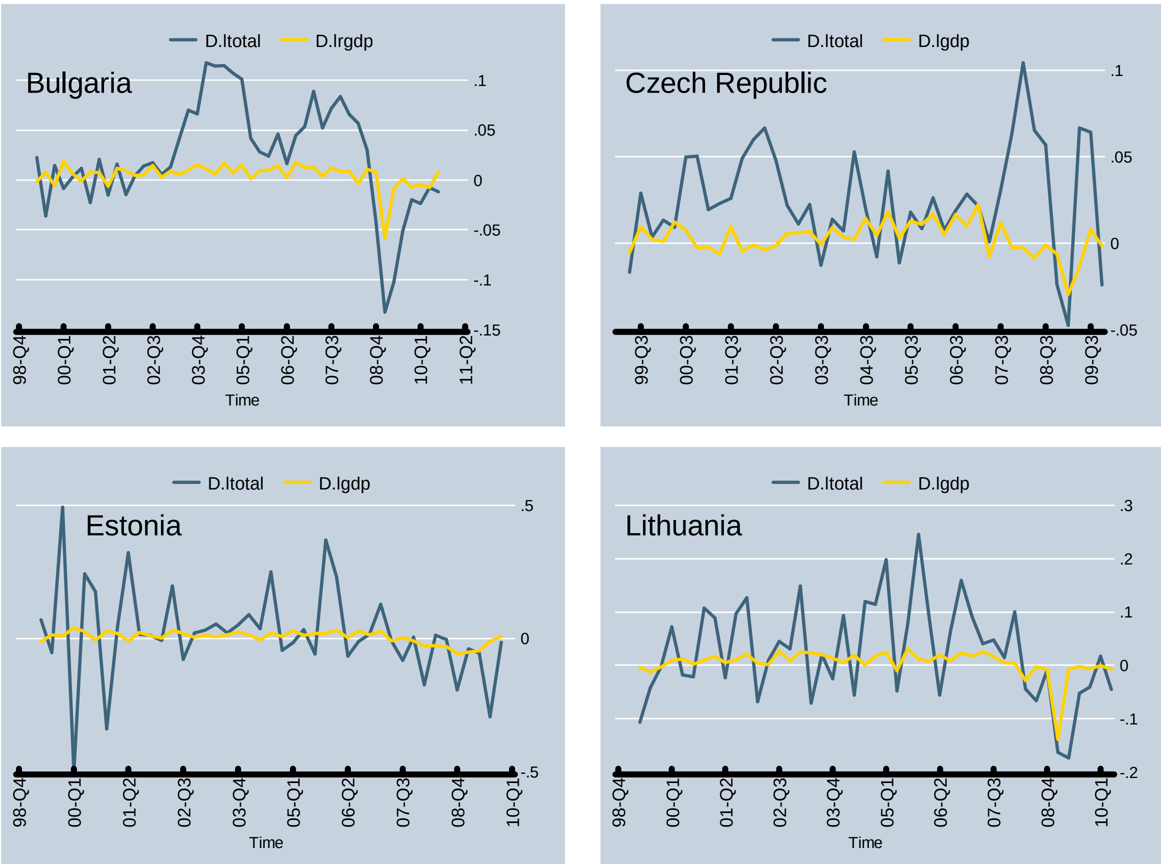
	Coef.	Std. Err.	z	P>z	[95% Conf. Interval]	
ltotal						
LD.	0.0152	0.146	0.1	0.917	-0.2702	0.3006
L2D.	-0.0106	0.131	-0.08	0.935	-0.2668	0.2456
lhicp						
LD.	-0.4906	0.420	-1.17	0.242	-1.3132	0.3320
L2D.	-0.6638	0.440	-1.51	0.131	-1.5258	0.1982
un						
LD.	-0.0123	0.010	-1.23	0.22	-0.0320	0.0074
L2D.	**-0.0269** **	0.011	-2.51	0.012	-0.0479	-0.0059
_cons	0.0498	0.014	3.62	0	0.0228	0.0768

Table D3) Results of Granger causality analysis

VECM		VECM		VAR		VAR	
BG		**CZ**		**EE**		**LT**	
omitted	p-value	omitted	p-value	omitted	p-value	omitted	p-value
HICP=>HP	0.173	GDP=>HP	**0.002**	GDP=>HP	**0**	HICP=>HP	**0.088**
LABF=>HP	**0.095**	UN=>HP	0.116	DIRINV=>HP	**0.005**	UN=>HP	**0.004**
MORTG=>HP	0.681						
L2HICP=>HP	0.0996	HP=>UN	**0.006**			HP=>UN	**0**
L2LABF=>HP	**0.03**						

This table contains results for all variables explaining house prices, and those variables whose past values turned out to affect house prices. The p-value at HICP=>HP marks the power of HICP in explaining HP. H_0 is that the former variable has no significant power in explaining the latter, i.e. A does not "Granger cause" B.

Figure D4) Dynamics of real house prices and real GDP



The graphs show quarter-on-quarter percentage changes in the variables.